

ISCRAM 2024
MÜNSTER, GERMANY

CONFERENCE PROCEEDINGS

21st International Conference on Information Systems for Crisis Response and Management

Münster, Germany, May 2024

Editors
Berthold Penkert
Bernd Hellingrath
Monika Rode
Adam Widera
Michael Middelhoff
Kees Boersma
Matthias Kalthöner

ISBN 978-3-00-079769-9

ISSN 2411-3387

ERCIS
Universität Münster
ERCIS C³M
IdF Institut der Feuerwehr NRW
STADT MÜNSTER

Table of Contents

ISCRAM 2024: Embracing the Crisis Management Lifecycle i
Program Committee iii
Organizing Committee iv
Sponsors v
Workshops vi
Project Symposium xiv
Track Chairs xxiv
Reviewers xxvi
ISCRAM 2024 Best Paper Awards xxix
Practitioner Track xxxi
Command & Control Studies 1
WiP Papers
Institutional amnesia during large crisis: Effects of poor documentation; Erik A.M. Borglund[1]; Kjell Lagerström[1]; [1]Mid Sweden University 2
Civil Protection in Europe: Towards a Unified Command System? Update on knowledge-sharing work and future actions; Jean Paul Monet[1]; Anaïs Saint-Jonsson[2]; Sergio Pirone[3]; Marc Castelnou Ribau[4]; Stephane Poyau[5]; Marc Dumas[1]; Miguel Cruz[6]; Tony Johansson[7]; Christoph Lamers[8]; Christos Lampris[9]; Pierre Schaller[10]; [1]Bouches-du-Rhône Fire Dpt; [2]Aix-Marseille University; [3]Regional Forest FireFighters Corp, Piedmont; [4]Catalonia Fire & Rescue Service; [5]Landes Fire Dpt, [6]Emergency and Civil Protection National Directorate, Portugal; [7]Södertörn Fire & Rescue Service; [8]Joint Competence Centre for Civil Protection NRW; [9]National Fire Service, Greece; [10]National Fire Service, France 10
Unified command? Preliminary findings from a situation awareness experiment; Jeroen Wolbers[1]; [1]LeidenUniversity 26
Collaborative Evaluation: Mapping Socio-Technical Patterns in an Emergency Exercise; Kjetil Rustenberg[1]; Kristine Steen-Tveit[1]; Bjørn Erik Munkvold[1]; Jaziar Radianti[1]; [1]Center for Integrated Emergency Management, University of Agder 35
Geographic Information Science (GIS) for Crisis Management 48
CoRe Papers
QuakeSet: A Dataset and Low-Resource Models to Monitor Earthquakes through Sentinel-1; Daniele Rege Cambrin[1]; Paolo Garza[1]; [1]Politecnico Di Torino 49
Volunteers in Crisis and Emergency Management 61
CoRe Papers
Managing volunteers in disaster situations - an overview of models for decision-makers; Florentina Hager[1]; Melanie Reuter-Oppermann[1]; [1]University of Twente 62
Improving emergency and disaster response management through bidirectional communication between citizens and decision-makers – A field test on the coordination of spontaneous volunteers; Michael Holzhüter[1]; Hans Betke[1]; Alina Fröhlich[1]; Michael Klafft[1,2]; [1]Fraunhofer Institute for Open Communication Systems; [2]Jade Hochshule 79
WiP Papers
The coordination and leadership in first aid ad-hoc immediate responder groups: a work-in-progress; Wilhelm Brodin[1]; Erik Prytz[1]; Carl-Oscar Jonson[1]; [1]Linköping University 97

Selection of rural villages as volunteer centers for emergency response; Niki Matinrad[1]; Tobias A. Granberg[1]; [1]Linköping University 108

Civilian-Military Integration in Ukrainian Defense Supply Chain; Myroslava Kushnir[1]; Anna Nagurney[2]; Renata Konrad[3]; [1]Ukrainian Catholic University; [2]UMass Amherst; [3]Worcester Polytechnic Institute 119

Usability and Universal Design of ICT for Emergency Management 131

CoRe Papers

Contextualizing Effective Use of Health Information Systems in Humanitarian Setting: The role of Transparent Interaction; Marta Maria Vila-Pozo[1]; [1]University of Oslo 132

WiP Papers

Avoiding "metal astray": experience evaluation of virtual reality training for melting furnace operators.; Jo Dugstad Wake[1]; Frode Guribye[2]; [1]NORCE; [2]University of Bergen 145

First Evaluation of the Usability and Usefulness of the PROACTIVE App; Laura Petersen[1]; Natasha McCrone[1]; Grigore Havârneanu[1]; Garik Markarian[1]; [1]UIC; [2]Rinisoft; [1]UIC; [2]Rinisoft..... 156

Integrated Wildfire Risk Management 166

CoRe Papers

Uncertainty in Forest Fire Management; Nikola Tietze[1]; Thomas Kox[1]; Till Büser[1]; Lars Gerhold[1]; [1]Freie University Berlin; [2]Weizenbaum Institute; [2]Weizenbaum Institute; [3];Technische Universitaet Braunschweig 167

WiP Papers

Networks for Integrated Wildfire Management: A Social Network Analysis Approach; Diogo Vallim[1]; Amalia Giannakopoulou[1]; Anna Coco Schmid[1]; David Struthers[1]; Efthymios Altsitsiadis[1]; Isabel Froes[1]; [1]Copenhagen Business School 176

Wildfire Evacuation: Social Resilience and Local Community Response..... 192

CoRe Papers

Evacuation modelling for community resilience: An Australian perspective; Dhirendra Singh[1]; Hanna Grzybowska[1]; Leorey Marquez[1]; Vincent Lemiale[1]; [1]CSIRO's Data61 193

Wildfire Evacuation Challenges in Multi-ethnic Communities: An Information Communication Perspective; Meng Duo[1]; Jun Hu[1]; Anying Chen[2]; Zhetao Fang[3]; [1]Beijing Normal University; [2]Jinan University; [3]China People's Police University..... 206

IT Solutions for Crisis Management..... 216

CoRe Papers

Drone Swarm for Post-Wildfire Hot Spot Detection: Technology Assessment and PoC Demonstrator; Micha Sende[1]; Christian Raffelsberger[1]; Samira Hayat[1]; Armin Köfler[2]; Alexander Almer[2]; [1]Lakeside Labs; [2]JOANNEUM RESEARCH..... 217

WiP Papers

Chatbot Based Public Sensoring to improve Situational Awareness; Hans Betke[1]; Thomas Kox[2]; Sascha Peitzsch[1]; Johannes Boldt[3]; Dustin Reimann[1]; [1]Fraunhofer FOKUS; [2]Weizenbaum Institute; [3]Martin Luther University Halle-Wittenberg 229

Modular Multi-Platform Interface to Enhance the Situational Awareness of the First Responders; Xabier Oregui[1]; Izar Azpiroz[1]; Veronica Ruiz[2]; Blanca Larraga[2]; Alvaro Gutierrez[2]; Igor G. Olaizola[1]; [1]Vicomtech

Foundation BRTA; [2]Universidad Politechnica de Madrid 242

Long-Range Human Detection in Drone Camera Images; Joris Heemskerk[1]; Tina Mioch[1]; Henry Maathuis[1]; Huib Aldewereld[1]; [1]University of Applied Sciences Utrecht 253

From Data to Action: A Graph-Based Approach for Decision Support in Civil Protection Operations Planning; Sabine Janzen[1]; Natalie Gdanitz[1]; Merlit Kirchhöfer[2]; Tobias Spanke[2]; Wolfgang Maass[1]; [1]German Research Center for AI, Saarbruecken; [2]Federal Agency for Technical Relief, Bonn 266

Adapting PLUM: Earthquake Early Warning with Node-Level Processing in New Zealand; Chanthujan Chandrakumar[1]; Marion Lara Tan[1]; Caroline Holden[2]; Max T. Stephens[3]; Raj Prasanna[1]; [1]Joint Centre for Disaster Research, Massey University; [2]SeismoCity Ltd.; [3]University of Auckland 275

A Study on The Frequency Characteristics of Typhoon Landing in Guangdong, China Based on Machine Learning Methods; Chenning Pan[1]; Xiaoyong Ni[1]; Ruoxi Lai[1]; Dafeng Ma[1]; Xifan Shen[1]; [1]Beijing Normal University 289

Improved satellite-based emergency mapping through automated triggering of processes; Martin Mühlbauer[1]; Monika Friedemann[1]; Johanna Roll[1]; Torsten Riedlinger[1]; Fabian Henkel[1]; Lucas Angermann[1]; Mathias Böck[1]; Tobias Kaminski[1]; Kayla Barginda[1]; [1]German Aerospace Center 305

SILVANUS Dashboard for Wildfire Management; Dominika Grunwald[1]; Marcin Wawrzyniak[1]; Marcin Przybyszewski[1]; Emil Gatial[2]; Zoltan Balogh[2]; Sepideh Hassankhani-Dolatabadi[2]; [1]ITTI Sp.zo.o., Poznań, Poland; [2]Slovak Academy of Sciences 314

Risk Communication and Community Engagement 327

CoRe Papers

Correlation Analysis of the Terrorist Attacks and Social Attributes in The Gulf States; Ting Liu[1]; Xiaoyong Ni[1]; Miaohan Shu[1]; Dafeng Ma[1]; Jiawei Sun[1]; [1]Beijing Normal University 328

Understanding Crisis Preparedness: Insights from Personal Values, Beliefs, Social Norms, and Personal Norms; Felix Burger[1]; Jasmin Haunschild[1]; Christian Reuter[1]; [1]Technical University of Darmstadt 345

WiP Papers

Engaging the Public in Forest Fire Awareness through a Mobile Application; Emil Gatial[1]; Zoltán Balogh[1]; Sepideh Hassankhani Dolatabadi[1]; Mariana Soledad Molina[2]; Timo Kasig[2]; Kayvan Yousefi Mojir[3]; Marcin Wawrzyniak[4]; Dominika Grunwald[4]; Marcin Przybyszewski[4]; [1]Slovak Academy of Sciences; [2]Massive Dynamic, Sweden; [3]University of Borås, Sweden; [4]ITTI Sp. z o.o. , Poland 360

Towards a Taxonomy for Conversational Agents in Disaster Management; Johannes Göbel[1]; Hans Betke[2]; Stefan Sackmann[1]; [1]Martin-Luther-University Halle-Wittenberg; [2]Fraunhofer FOKUS 373

Tweeting Through the Flood: Application of BERT Topic Modeling for a Comparative Flood Communication Analysis; Christin Salley[1]; Nathan Fox[1]; Alyssa Schubert[1]; [1]Michigan Institute for Data Science 384

Broadcasting Resilience: Developing Wildfire Readiness Through Student-Driven Radio Programs; Jose Azevedo[1]; Susana Neves[1]; Ivone Neiva Santos[1]; [1]University of Porto 395

Social Media for Crisis Management 401

CoRe Papers

Semantically Enriched Cross-Lingual Sentence Embeddings for Crisis-related Social Media Texts; Rabindra Lamsal[1]; Maria Rodriguez Read[1]; Shanika Karunasekera[1]; [1]The University of Melbourne 402

Image-text crisis tweet categorization: a caption-based approach; Badreddine Farah[1]; Guillaume

Cleuziou[1]; Cécile Gracianne[2]; Adel Hafiane[1]; Anaïs Halftermeyer[1]; Raphaël Canals[1]; [1]University of Orleans; [2]BRGM France 414

Enhancing Emergency Post Classification through Image Information Amplification via Large Language Models; Pablo Giaccaglia[1]; Carlo Alberto Bono[1]; Barbara Pernicis[1]; [1]Politecnico di Milano 426

Monitoring Critical Infrastructure Facilities During Disasters Using Large Language Models; Abdul Wahab Ziaullah[1]; Ferda Ofli[1]; Muhammad Imran[1]; [1]Hamad Bin Khalifa University, Doha, Qatar 439

Empowering Crisis Response Efforts: A Novel Approach to Geolocating Social Media Images for Enhanced Situational Awareness; Hafiz Budi Firmansyah[1]; Jose Luis Fernandez-Marquez[1]; Mehmet Oguz Mulayim[2]; Jorge Gomes[3]; Joao Ribeiro[3]; Valerio Lorini[4]; [1]University of Geneva; [2]Spanish National Research Council; [3]VOST Europe Santarem, Portugal; [4]Joint Research Center, European Commission Ispra, Italy 453

WiP Papers

Social Media Analysis in Sudden Onset Disasters and its Usefulness for Decision Makers - A Scoping Review; Francesca Müller[1]; Sylvia Bach[1]; Frank Fiedrich [1]; [1]University of Wuppertal 467

Utilizing Social Media and Crowdsourcing for Crisis Management and Communication: proposing an Integrative Framework; Kees Boersma[1]; Nathan Clark[1]; Robert Larruina[1]; Emmeline Roeloffs[1] 482

Effectiveness of Emergency Response to Typhoons: Lessons Learned from Historical Typhoon Disasters in Zhuhai, China; Ting Liu[1]; Xiaoyong Ni[1]; Jiawei Sun[1]; Ziyi Peng[1]; [1]Beijing Normal University 493

Forced Displacement Situation Awareness with Large Language Models: Ukrainian Refugees in Poland Case Study; Brian Tomaszewski[1]; Nataliya Shakhovska[2]; Harley Emery[3]; Paweł Śniatała[4]; [1]Rochester Institute of Technology; [2]Lviv Polytechnic National University; [3]University of Maryland; [4]Poznan University of Technology 508

Crisis[2]Sum: An Exploratory Study on Disaster Summarization from Multiple Streams; Philipp Seeberger[1]; Korbinian Riedhammer[1]; [1]Technische Hochschule Nuernberg 523

Can User Characteristics Predict Norm Adherence on Social Media? Exploring User-Centric Misinformation Interventions; Shangde Gao[1]; Yan Wang[1]; [1]University of Florida 544

M-CATNAT: A Multimodal dataset toanalyze French tweets duringnatural disasters; Badreddine Farah[1]; Omar El Bachyr[1]; Guillaume Cleuziou[1]; Anaïs Lefeuvre-Halftermeyer[1]; Cécile Gracianne[2]; Samuel Auclair[2]; Adel Hafiane[1]; Raphael Canals[1]; [1]University of Orleans; [2]BRGM France 555

Impact-driven Emergency and Crisis Management 567

CoRe Papers

State of Knowledge Management among German fire fighters; Andreas Schultz[1]; Fabian Dotzki[1]; Iryna Mozgova[1]; [1]Paderborn University 568

WiP Papers

Cascading supply chain disruption: impact on water industry during COVID-19; Lijo John[1]; Wojciech Piotrowicz[1]; Aino Ruggiero[2]; [1]Birla Institute of Technology Pilani-Dubai Campus; [2]Hanken School of Economics, Helsinki 583

Assessing Infrastructure Resilience in Small Islands: Methodology Proposal; Fabiane A. Acordes[1]; Sylvia Bach[1]; Frank Fiedrich[1]; Frank Sill Torres[2]; [1]University of Wuppertal; [2]German Aerospace Center 591

A Study on The Heterogeneity of Tropical Cyclone Vulnerability Functions in Different Provinces along The Coastal Areas in China; Chenning Pan[1]; Xiaoyong Ni[1]; Xifan Shen[1]; Ziyi Peng[1]; [1]Beijing Normal University 602

Crisis Response and Management of Public Administrations: Forgotten Tales; Yannic Schulte[1]; Malte Schönefeld[1]; Patricia M. Schütte[1]; Frank Fiedrich[1]; [1]University of Wuppertal 617

Uncovering Impact of Innovation: Continuous Stakeholder Engagement through Scenario-based Systems Engineering; Jens Pottebaum[1]; Marcel Ebel[1]; Iris Gräßler[1]; [1]Paderborn University 626

Disaster Public Health and Healthcare Informatics 638

CoRe Papers

Accounting for systemic risk in analysis of crisis scenarios; Jose J. Gonzalez[1]; Reem Abbas[2]; Sigurd Paulsen[3]; [1]University of Agder, Norway; [2]Auckland University of Technology; [3]Dept. of Organization, Crisis Management and Information Security, Kristiansand municipality, Norway..... 639

Perceiving and Behaving in a Crisis: Developing a Multi-Functional Crisis Information Platform for Psychosocial Situations (CIP-PS); Sebastian Sterl[1]; Andreas Billig[2]; Franck William Taffo[2]; David Svoboda[2]; Lars Gerhold[1]; [1]Technische Universitaet Braunschweig; [2]Fraunhofer Institute for Open Communication Systems 654

Designing Information Flow with Diverse Community Health Workers to Improve Public Health Situational Awareness During COVID-19; Temitayo Okusanya[1]; Jane Gresser[2]; Michael Stevenson[3]; Ajay Kumar[4]; Zeno Franco[1]; Md Romael Haque[2]; [1]Medical College of Wisconsin; [2]Marquette University; [3]University of Wisconsin-Madison; [4]University of Muenster 672

Psychosocial Aspects of the 2021 Flood Disaster in Anonymous District, Germany: Strains, Resources, and the Role of Social Media; Francesca Müller[1]; Marvin Kubitza[1]; Frank Fiedrich[1]; [1]University of Wuppertal 687

Visions for Future Crisis Management 699

WiP Papers

Technology Enhanced Actual Situation Awareness Model for Dog-Handlers working with Search and Rescue Dogs; Stella Polikarpus[1]; Kristine Steen-Tveit[2]; Ana María Aldea Reyes[3]; Marta Burgos González[3]; Susana Izquierdo Funcia[3]; Graham Edgar[4]; [1]Estonian Academy of Security Sciences; [2]University of Agder; [3]Spanish School of Rescue and Detection with Dogs; [4]University of Gloucestshire 700

Exploring the roles and challenges of disruptive technologies for emergency management in Aotearoa New Zealand; Lukas Kroher[1]; Marion Lara Tan[2]; Raj Prasanna[2]; [1]Otto-Friedrich University Bamberg; [2]Massey University 709

Vision for Emergency Management: Conceptualizing Mission-Critical Ecosystems; Jaziar Radianti[1]; [1]University of Agder 721

Bridging Gaps, Building Futures: How DIREKTION Innovates for Global Disaster Resilience; Juliane Schlierkamp[1]; Eileen Murphy[2]; Louise Vos[3]; [1]Fraunhofer INT; [2]Trilateral Research; [3]TNO..... 734

Pre-disaster Training for Firefighters Using Augmented Reality: Hologram Characteristics; Stella Polikarpus[1]; Reet Kasepalu[1]; Kärt Reitel[1]; [1]Estonian Academy of Security Sciences 744

Information Systems in Humanitarian Operations 757

WiP Papers

Navigating the Digital Technology Behavior for the Distribution of Humanitarian Cash-Based Assistance: Aid Recipients Experience; Amin Maghsoudi[1]; Abdelsalam Hamid Abakar[1,2]; [1]Hanken School of Economics; [2]A'Sharqiyah University 758

Analytical Modeling and Simulation 767



CoRe Papers

Possibility of Suppressing Fire-spread in the Event of a Large Earthquake by Greening and Demolishing Vacant Houses; Toshihiro Osaragi[1]; Asuka Harada[1]; Takuya Oki[1]; Yoho Sakamoto[1]; [1]Tokyo Institute of Technology 767

Enabling Information Transmission in Low-Throughput Wireless Channels for Aerial Disaster Monitoring Systems; Julian Zobel[1]; Ralf Steinmetz[1]; [1]Technical University of Darmstadt..... 785

WiP Papers

Location and allocation of primary and backup shelters in transboundary disasters; Oscar Rodriguez-Espindola[1]; Pavel Albores[1]; Dolores Edwiges Luna-Reyes[2]; Juan Antonio Diaz-Garcia[2]; Sergio Picazo-Vela[2]; [1]Aston Business School; [2]Universidad de las Americas Puebla..... 801

Using Optimization Modelling to Determine Locations for Aerial Firefighting Resources Based on Wildfire Risk Data; Viktor Sköld Gustafsson[1]; Tobias Andersson Granberg[1]; Martin Waldemarsson[1]; [1]Linköping University 813

Towards a repository for datasets of mass casualty incidents; Florentina Hager[1]; Melanie Reuter-Oppermann[1]; [1]University of Twente 822

Cascading Effects of Critical Infrastructures in a Flood Scenario: A Case Study in the City of Cologne; Moritz Schneider[1]; Peter Priesmeier[2]; Alexander Fekete[2]; Daniel Lichte[1]; Frank Fiedrich[3]; [1]German Aerospace Center; [2]Cologne University of Applied Sciences; [3]University of Wuppertal..... 834

Homophily in opinion networks affects collective risk perception in heterogeneous populations; Francesca Giardini[1]; Daniele Vilone[2]; Lorenzo Zino[3]; Ming Cao[1]; [1]University of Groningen, The Netherlands; [2]National Research Council, Rome, Italy; [3]Politechnico di Torino 843

Sensitivity Analysis of a Road Clearing and Relief Supplies Distribution Model for Vancouver Island in the Event of a Cascadia Megathrust Earthquake; Ronald Pelot[1]; Pranitha Vattoni[1]; Luana Souza Almeida[1]; Floris Goerlandt[1]; [1]Dalhousie University, Halifax, Canada 852

Open Track 870

CoRe Papers

R²DN – Integrating Databases about Natural Disasters; Jose Verissimo Pereira[1]; Marcos Borges[1]; [1]Universidade Federal do Rio de Janeiro..... 871

The Role of Digitalization During Supply Chain Disruptions: Lessons from the Experiences of Family-Owned Austrian Firms During the COVID-19 Pandemic; Deepak Khazanchi[1]; Reinhard Bernsteiner[2]; [1]University of Nebraska Omaha; [2]Management Center Innsbruck Austria 888

Comparative Analysis of Strategies for National Food Stockpiling: A Case Study of Germany and Switzerland; Katharina Eberhardt[1]; Simon Stieler[1]; Florian Klaus Kaiser[1]; Frank Schultmann[1]; [1]Karlsruhe Institute of Technology..... 898

Empirical Investigation of Values Affected by AI Systems for Fire Services; Tina Mioch[1,2]; Huib Aldewereld[1]; Mark A. Neerincx[2]; [1]Delft University of Technology; [2]University of Applied Sciences Utrecht 912

WiP Papers

Introducing WALKER – Safeguarding Cultural Heritage through Evacuation Run Maps; Tobias Hellmund[1]; Jürgen Moßgraber[1]; Jürgen Reuter[1]; Susanne Rißmann[2]; Hermann Bayer[3]; Stefan Deschermeier[3]; [1]Fraunhofer IOSB; [2]Bayerische Verwaltung der staatlichen Schlösser; [3]Kreisfeuerwehrverband München e.V. 924

The Utility of Community Resilience Models for the Prevention of Online Violence; Dante Della Vella[1];

Hoang Long Nguyen[1]; Rajendra Akerkar[1]; [1]Vestlandsforsking 931

Scoping Review: Digitalisation in the organisational culture of safety and security agencies; Thomas Kox[1]; Patricia Schütte[2]; Sebastian Rohwedder[1]; Till Büser[1]; [1]Weizenbaum Institute; [2]Bergische Universitaet Wuppertal 942

Human Values for Responsible Decision-Support for Fire Services; Tina Mioch[1,2]; Huib Aldewereld[1]; Mark A. Neerincx[2]; [1]Delft University of Technology; [2]University of Applied Sciences Utrecht 952

Inter-professional simulator training for reliable ship-helicopter interactions; Thomas Lübcke[1]; Hendrik Wilhelm[2]; Oda Schliebusch-Jacob[1]; Tim Kreter[2]; Tanja Martini[3]; Torsten Gerlach[3]; [1]German Maritime Search and Rescue Service; [2]Witten/Herdecke University; [3]German Aerospace Center 961

User-tailored visualization of simulation and sensor data for efficient crisis management; Nils Winter[1,4]; Michael Monteforte[1,3]; Fabian Müller[1,2]; Ines Rohrbach[1,3]; Philipp Hertweck[1,4]; Till Martini[1,2]; [1]Fraunhofer SIRIOS, Berlin, Germany; [2]Fraunhofer EMI, Freiburg, Germany; [3]Fraunhofe IVI, Dresden, Germany; [4]Fraunhofer IOSB, Karlsruhe, Germany 972

Strategies for Crisis-Responsive Governance: Automated Anomaly Identification in Public Services; Hakan Unveren[1]; Ayesh A. Lehyeh[1]; Duygu Pamukcu[2]; Christopher W. Zobel[3]; [1]University of Vermont; [2]HEC Montreal; [3]Virginia Tech 981

Scenario planning to enable foresight in crisis management; Marte Luesink[1,2]; Jeroen Wolbers[1]; Menno van Duin[1]; Sanneke Kuipers[2]; [1]Nederlands Instituut Publieke Veiligheid; [2]LeidenUniversity 995

One System to Connect them All - A Core System for Realizing Integrated Command and Control Research; Björn Johansson[1]; Jesper Tordenlid[2]; Jonas Lundberg[1]; Jens Alfredson[3]; [1]Linköping University; [2]Combitech AB; [3]Saab AB, Aeronautics 1004

Towards training network resilience to maintain disaster recovery expertise; Paul Savignac[1]; Eva Petitdemange[1]; Christopher Zobel[2]; Nicolas Cellier[1]; Robert Weiss[2]; Matthieu Lauras[1]; [1]Centre de Genie Industriel, IMT MINES ALBI, Universite de Toulouse; [2]Virginia Tech 1015

Barriers against data sharing for Disaster Risk Reduction (DRR) at borders; Anouck Adrot[1]; Lola Brunet[1]; Henri Isaac[1]; [1]Université Paris-Dauphine, Université PSL, CNRS, DRM 1025

End-user perspective on the results and impact of European projects: a case of failure to innovate; Maaike van de Vorst[1]; Kenny Meesters[2]; Bert Brugghemans[3]; [1]APB Campus Vesta; [2]Tilburg University; [3]Brandweerzone Antwerpen 1037

Knowledge on demand: the future of decision support systems?; Yohann Chasseray[1]; Aurélie Congès[1]; Audrey Fertier[1]; Anne-Marie Barthe-Delanoë[1]; [1]IMT Mines Albi, Universite de Toulouse.. 1050

Future AI in crisis management: Proposing a bio-inspired, neuro-symbolic architecture; Aurélie Montarnal[1]; Frédérick Bénaben[1]; Rob Grace[2]; Grégoire Borst[3]; Paul Gaborit[1]; [1]IMT Mines Albi; [2]Texas Tech University; [3]LaPsyDE - UMR CNRS 8240, Université Paris Cite 1061

Envisioning the Future of Emergency Management focused on Artificial Intelligence; Ana-Gabriela Núñez[1]; Patricia Quiroz-Palma[2]; M.Carmen Penadés[3]; Otto Parra[1]; Oscar Alvear[1]; [1]Universidad de Cuenca; [2]Universidad Laica Eloy Alfaro de Manabi; [3]Universitat Politecnica de Valencia, Spain 1072

Evacuation Social Support Among New Orleans Households: Implications for Evacuation Management Systems; Rob Grace[1]; Hyeong Suk Na[2]; [1]Texas Tech University; [2]University of Missouri 1084

WUI-PEM: Wildfire Phased Zone Evacuation Methodology; Hui Xie[1]; Michael Kinsey[1]; Steve Gwynne[1]; Max

Kinateder[2]; Paul Geoerg[3]; Roy Thun[4]; [1]Movement Strategies; [2]NRC Canada; [3]Vereinigung zur Förderung des Deutschen Brandschutzes e.V., Germany; [4]GHD, USA 1096

It's a Machine Learning world: no future for Complex Event Processing?; Anne-Marie Barthe-Delanoë[1]; [1]IMT Mines Albi, University of Toulouse 1107

Posters 1119

Demonstrations 1151

ISCRAM 2024: Embracing the Crisis Management Lifecycle

The conference follows a pracademic approach and will be jointly hosted by the European Research Center for Information Systems (ERCIS) Competence Center for Crisis Management at the University of Münster and the State Fire Service Institute North Rhine-Westphalia, Germany. As seen and discussed in the ISCRAM community in recent years, the need of "pracademics" became essential to facilitate and further increase high-quality research on information systems in crisis management (CM) and response. Thus, it appeared as a logical consequence to join forces from both the practitioner and scientific communities to promote the ISCRAM mission. The "pracademic approach" is considered for all relevant conference elements, such as a balanced keynote lineup or practitioner-friendly contributions (e.g. practitioner talks instead of practitioner papers).

Besides, the conference theme takes a process-centric view on crisis management: "Embracing the Crisis Management Lifecycle". The theme emphasizes a holistic and integrated process view of crisis management to overcome isolated considerations of crisis management lifecycle phases and showcase how important well-structured and clearly defined roles, tasks and processes are to supporting communication and collaboration between involved actors.

The ISCRAM 2024 Proceedings include two types of papers: Completed Research (CoRe) and Work in Progress (WiP) Papers. CoRe papers are longer in length and describe completed research studies. Only CoRe papers were candidates for the Best Research Paper or Best Student Paper awards. WiP papers are shorter and meant to describe preliminary or more focused research findings. Both paper types (CoRe and WiP) underwent a rigorous peer-review process to determine the papers included in these proceedings and archived in the ISCRAM digital library.

In total, 31 CoRe papers were submitted and 27 were accepted. Two out of the four rejected papers got recommended for a resubmission as WiP. The overall acceptance rate for the CoRe papers at the conference was 87%. Besides, 96 WiP papers were submitted and 66 got accepted. The overall acceptance rate for the WiP papers at the conference was 69%.

Besides, practitioners' contributions are documented as abstracts of the new practitioner talks track as well as project symposium abstracts. Practitioner talks are peer reviewed to high ISCRAM standards and their presentations were similar in length to CoRe paper presentations. In total, seven practitioner talks were submitted and accepted.

The program is completed with eight workshops, ten posters and two demonstrations. On top of that, a Project Research Symposium under the topic: "Fostering collaboration: Bridging Academia and Practice through ongoing Projects and Opportunities for Future Research". The project symposium is designed to foster interactions between academia and practitioners by showcasing funded research projects, both ongoing and completed, encouraging discussions about future research ideas and finding synergies. In total, nine projects were accepted. The project symposium presentations were limited to seven minutes each, followed by a world-café format.

Five criteria (relevance, significance, originality, validity, and clarity) were used to review all types of ISCRAM papers by at least two regular and one meta review. Papers were submitted to one of fifteen tracks, which included:

- Command & Control Studies

- Geographic Information Science (GIS) for Crisis Management
- Volunteers in Crisis and Emergency Management
- Usability and Universal Design of ICT for Emergency Management
- Integrated Wildfire Risk Management
- Wildfire Evacuation: Social Resilience and Local Community Response
- IT Solutions for Crisis Management
- Risk Communication and Community Engagement
- Social Media for Crisis Management
- Impact-driven Emergency and Crisis Management
- Disaster Public Health & Healthcare Informatics
- Visions for Future Crisis Management
- Information Systems in Humanitarian Operations
- Analytical Modeling and Simulation
- Open Track

All authors in these proceedings have agreed to the ISCRAM 2024 Proceedings Copyright Agreement. This copyright agreement and use license state that an author retains copyright. It also permits any user – for any non-commercial purpose, including unlimited classroom and distance learning use – to download, print out, extract, archive, distribute and make derivative works of an article published in the ISCRAM 2024 Proceedings, if appropriate credit is given to the authors, the source of the work, and all derivative works are placed under the same license.

We would like to recognize the efforts of the Organizing Committee and Program Committee (track chairs and reviewers) for making the ISCRAM 2024 conference possible. We also thank the ISCRAM board, our partners, and sponsors for their significant support of the conference. Lastly, we sincerely thank the authors and speakers who shared their work at ISCRAM 2024. We thank you all for your service to the ISCRAM community over the past year and look forward to the next ISCRAM world conference in Halifax, Canada.

Program Committee

Program Chairs

Dr. Michael Middelhoff, University of Münster, Germany
Prof. Dr. Ir. Kees Boersma, Vrije Universiteit Amsterdam, Netherlands
Matthias Kalthöner, State Fire Service Institute North Rhine-Westphalia, Germany

Workshop Chairs

Dr. Christoph Lamers, Joint Competence Center for Civil Protection, Germany
Johannes Ponge, M.Sc., University of Münster, Germany
Prof. Jaziar Radianti, University of Agder, Norway

Practitioner Talks Chairs

Nicola Rupp, Deutsches Rettungsrobotik Zentrum e.V., Germany
Bart van Leeuwen, Netage B.V. and Amsterdam-Amstelland FD, Netherlands
Dr. Christoph Lamers, Joint Competence Center for Civil Protection, Germany

Doctoral Colloquium Chairs

Prof. Dr. Melanie Reuter-Oppermann, University of Twente, Netherlands
Prof. Dr. Stefan Klein, University of Münster, Germany
Prof. Dr. Mark Haselkorn, University of Washington, USA

Demonstration and Poster Chairs

Dr. Sandra Wissing, FH Münster, Germany
Dr. Monika Rode, State Fire Service Institute North Rhine-Westphalia
Prof. Dr. Apoorva Chauhan, Concordia University of Edmonton, Canada

Project Symposium

Dr. Chiara Fonio, ICONS Innovation Strategies, Italy
Dr. Alexander Rösner, Federal Agency for Technical Relief, Germany
Dr. Monika Rode, State Fire Service Institute North Rhine-Westphalia, Germany
Dr. Adam Widera, University of Münster, Germany

Awards Committee Chair

Filippo SanFilippo, University of Agder, Norway

Organizing Committee

Conference Chairs

Berthold Penkert, State Fire Service Institute North Rhine-Westphalia, Germany
Prof. Dr.-Ing. Bernd Hellingrath, University of Münster, Germany
Dr. Monika Rode, State Fire Service Institute North Rhine-Westphalia, Germany
Dr. Adam Widera, University of Münster, Germany

Publicity and Communication Chairs

Stephanie Albiero, MA Int.Sec., United Nations Office for Project Services, USA
Sebastian Henke, M.Sc., University of Münster, Germany

Proceedings Chair

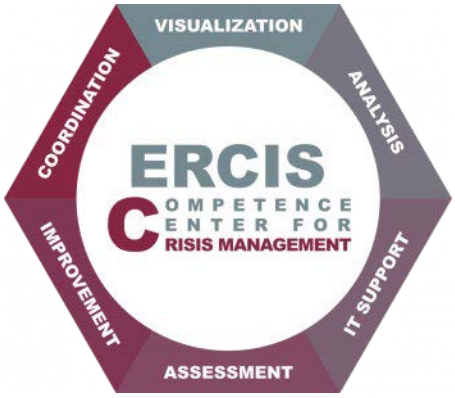
Anirudh Ravi, University of Münster, Germany

Conference Local Management and Logistics

Carmen Sicking, University of Münster, Germany

Sponsors


ERCIS


VISUALIZATION
ANALYSIS
IT SUPPORT
ASSESSMENT
IMPROVEMENT
COORDINATION
ERCIS
COMPETENCE CENTER FOR RISIS MANAGEMENT


Universität
Münster


IdF
Institut der Feuerwehr NRW


STADT MÜNSTER

Workshops

1. Workshop: Remote Crowdsourcing and Participatory Local Mapping for Disaster Management

Oliver Fritz, HeiGIT gGmbH
Anne Schauß, HeiGIT gGmbH
Marcel Reinmuth, HeiGIT gGmbH

The workshop, led by HeiGIT, focuses on addressing the challenges of accessing and capturing spatial insights throughout the disaster management cycle. Participants will explore the Sketch Map Tool and CrowdMap, open-source tools designed to collect and map local knowledge and interpret imagery through participatory mapping approaches. Through hands-on sessions and discussions, attendees will learn how these tools can engage local communities and remote volunteers to provide valuable data, contributing to more effective disaster management practices. The workshop provides a platform for sharing experiences and discussing opportunities and challenges in integrating crowdsourcing and participatory mapping into disaster management strategies.

2. Workshop: Devising Assessment and Mitigation of Systemic Risks with Strategy Mapping with a Participant Group

Jose J. Gonzalez, University of Agder
Sigurd Paulsen, Crisis Manager at Kristiansand Municipality, Norway

In our interconnected society, risks create cascading effects across sectors, forming a complex system with interdependent outcomes. Involving practitioners and experts is crucial for effective systemic risk assessment and management. A strategy mapping method, originally developed for complex projects in the 1990s, has been extended to address systemic risks in societal resilience and pandemics, notably applied to Covid-19.

3. Workshop: Using Bomb Shelters in the 21st Century?

Jaziar Radianti, University of Agder
Terje Gjøsæter, University of Agder
Stella Polikarpus, Estonian Academy of Security Sciences
Priit Laaniste, Estonian Academy of Security Sciences and Estonian Ministry of the Interior

The workshop examines the importance of utilizing bomb shelters in contemporary times, acknowledging insights from the conflict in Ukraine and the historical neglect of shelter maintenance since the Cold War. It delves into the potential of new technologies in warning the public about attacks while addressing associated security risks. Through discussions on information management, public training, and warning systems, participants aim to compile a thorough list of research topics pertinent to the modern utilization of bomb shelters, considering evolving societal expectations and the role of social media in crisis communication.

4. Workshop: Measuring Stress Levels in Critical Systems for Monitoring and Response

Michiel Rhoen, Veiligheidsregio Utrecht
Luuk Vink, Veiligheidsregio Utrecht

The Dutch Safety Regions utilize data streams to monitor stress levels and risk indicators across essential systems, enabling measurement of relative stress levels, real-time crisis response support, and long-term trend analysis. Emphasis is placed on both crisis response systems and

everyday societal continuity essentials such as telecommunications, power, sewers, and water management. The goal of the workshop is to present examples of data usage for stress measurement in critical systems across different time periods, solicit feedback for the development of a real-time risk monitor, and promote collaboration between first responders, data scientists, and experts in critical systems.

5. Workshop: Using Emergency Response Data to Connect Practice and Science

Amy Matser, Netherlands Institute for Public Safety
Guido Legemaate, Safety Region of Amsterdam – Amstelland
Vera Oosterveen, Netherlands Institute for Public Safety
Michiel Rhoen, Veiligheidsregio Utrecht (VRU, Safety Region Utrecht), Netherlands
Rob Peters, Utrecht University of Applied Scientific Research

The workshop explores the utilization of emergency response data for both practice and scientific research, bridging the gap between operational needs and academic inquiry. By aligning with the ISCRAM 2024 theme of embracing the crisis management lifecycle, the workshop aims to extend understanding beyond traditional boundaries. Drawing from over a decade of experience in the Netherlands, participants will delve into real projects, discussing success criteria and methods to meet scientific standards while upholding the quality of emergency response systems. Through interactive sessions, attendees will collaboratively develop research proposals focusing on key themes such as cross-border interoperability, event management, resilience, crisis communication, floods, and infectious agent spread.

6. Workshop: Cascading Effects in Complex Health Networks

Stefan Schauer, AIT Austrian Institute of Technology
Manuel Egger, AIT Austrian Institute of Technology

The workshop focuses on critical infrastructures and entities essential for maintaining societal functions, highlighting the interdependence and complexity of these systems, particularly in domains like healthcare. With increasing digitalization, cybersecurity becomes paramount, especially in sectors like healthcare, which are prime targets for adversaries. The workshop aims to provide a holistic understanding of cyber-physical systems and their interdependencies to enhance risk and resilience management in critical domains.

7. Workshop: Capability Gaps in Fire Fighting – how to proceed

Juliane Schlierkamp, Fraunhofer INT
Maike Overmeyer, Fraunhofer INT
Jean-Paul Monet, Bouches-du-Rhône Fire Department
David Martín, Pau Costa Foundation

The workshop, facilitated by the DIREKTION and FIRELOGUE projects, aims to foster collaboration and innovation in firefighting by bringing together practitioners and solution innovators. Designed for both industry experts and emergency responders, it provides a platform to discuss challenges, share insights, and explore solutions in wildfire risk management. Participants can expect lively discussions, knowledge exchange, and opportunities to shape the future of firefighting through collaboration and research-driven innovation.

8. Workshop: Netcentric Crisis Management – a Dutch approach and its challenges

Steven van Campen, Trimension
Willem Treurniet, Netherlands Institute for Public Safety – NIPV
Arthur van Lohuijzen, Trimension

The workshop, led by experts from the Netherlands Institute of Public Safety (NIPV) and Trimension, focuses on Netcentric Crisis Management (NCM) as a method to enhance information management and coordination during crises. Aimed at emergency responders, academics, researchers, and policymakers, it offers interactive sessions to discuss experiences, principles, implementation challenges, and future developments of NCM. Participants engage in cross-organizational knowledge exchange and exploration of innovative information management processes and tools to enrich NCM practices.

Netcentric Crisis Management
Insights gathered during a workshop on the ISCRAM2024 conference

How do you facilitate collaboration between involved actors during a crisis? How do you help them to overcome isolated considerations? How to effectively manage a complex network, stimulating shared situational awareness? All these questions underline that in many ways crisis management is network management. Over the past 20 years, emergency services, government, healthcare and critical infrastructure organisations in the Netherlands have co-created Netcentric Crisis Management (NCM); a method by which information management effectively supports decision making and coordination during crises.

During the ISCRAM2024 conference, NCM-experts of the Netherlands Institute of Public Safety (NIPV) and Trimension hosted a workshop in which the basic principles of NCM together with its impact were explained, the steps and challenges for implementation presented and the fits and non-fits of these were discussed in round-table sessions. The findings from both sessions are documented in this report and shared afterwards with the participants.

Set-up of the workshop

The workshop was aimed at emergency response practitioners, academics, researchers and policy makers with a specific interest in information management and netcentric operations to improve crisis management. The main aim was to discover further possibilities for NCM, obtain feedback from the international conference audience and discover state-of-the-art information management processes, systems and tools to further enrich NCM. The workshop consisted of two interactive round-table sessions of 90 minutes each. Participants were encouraged to attend both sessions and were mixed in smaller, multidisciplinary groups, thereby stimulating a cross-organisational exchange of knowledge, challenges and experiences.

The first session consisted of:

- An introduction to the workshops program, what NCM is, its history and basic principles, combined with practical uses of NCM during crisis situations.
- A first round-table session aimed at the questions: Do you recognise similar problems with information and crisis management in your region or sector? How do you cope with them? Would NCM fit your organisation's culture and processes?
- A short recapitulation of the most interesting results per table.
- The second session was focussed on implementing NCM, the challenges and potential developments and consisted of:

- A presentation on how NCM can be implemented in organisations and in networks from practical experience and the challenges and potential developments still ahead.
- Second round table session aimed at the questions: Do you recognise the presented implementation challenges? How could an implementation of NCM be successful in your country or organisation and which challenges do you foresee? What further developments or improvements do you suggest?
- A short recapitulation of the most interesting results per table.

Participants were split in two groups. The organisers provided canvases for both sessions, to document the findings in each round-table group.

Results from Session 1:
Question 1: Do you recognise the presented issues and challenges in information and crisis management in your own professional environment? If so, how?

Table 1	**Table 2**
Norway • Responsibilities: take responsibility vs. Handing over vs. Putting down • Covid-pandemic: Innovative risk method management with strict responsibilities; how emergency services are connected with each other: • Issues with data sharing due to national security concerns, therefore, specific mapping of information for chosen emergency service units. **Netherlands** • Train accident: Reaching out to relatives of casualties; who is responsible for accident?; privacy regulations are blocking sometimes • Limburg-south: Healthcare support & supporting infrastructures and including them into information sharing for measuring systems, detection systems for analysing **Italy** • (De-)centralised system for disaster management • Dealing with the amount of untrained volunteers **Germany** • Decentralised disaster management	1. Train for network centric -> high barrier (Netherlands) 2. No shared system (Norway + Catalonia) 3. Organisational boundaries (Norway + Catalonia) 4. Reluctance to share info (Belgium + Norway + Catalonia) • Passive = easier • No shared training + professional culture • 5. Hierarchy + professional culture (Catalonia) • Masculine culture 6. C&C tension between tactical and strategic level (Netherlands + Belgium + Norway) 7. Security + confidentiality (Netherlands + Belgium + Norway) • Different meaning of what a COP is

• Data protection law • • No army support	

Question 2: How could an implementation of NCM be successful in your country or organisation, and which challenges do you foresee?

Table 1	Table 2
• Strategy mapping1 • Exercises • IT-solutions • Trust building • Using social media	4. Trust building -> train together, swift trust -> see others act prof. -> legitimate by law -> expectation management 5. More women in emergency response + retaining women + train both masculine and feminine leadership styles 7. Wink wink nudge nudge; open up information in C&C -> case by case • Shared goal finding + prevent institutional crisis • Split between Command&Control and Info management processes

Question 3: Would NCM fit your culture and processes? If so, how?

Table 1	Table 2
• Networks: Bringing "emerging services" together with first responders • Sense-making: Every helping hand has to feel their support value • Common Operational Picture: IT solution(s) and merging them • For Command&Control: Single or multiple command posts	• Norway: Striving to get there • Belgium: First single agency + national agreement • Catalonia: Emergency services work together but reluctance at non-emergency services

Results from Session 2

Question 1: Do you recognise the presented implementation challenges?

Table 1	Table 2
• Belgium: Cultural issues o Languages o Organisations not talking to each other o People not talking to each other • Norway: Budget cuts • Germany: Reasonable service- blueprint for emergency service units; supply chain management similarities with multiple self-centric goals; issues with "global optimum"-goal reaching multi-facetted	• Cultural/political elements o Drivers/Barriers vary between cultures considerably • Limited time/resources to participate • Complexity of network collaboration o Roles regime with different types of knowledge o Need for power brokers, with the right connections and ability to execute o Identifying right mix of individuals

problems (legally, organisation, IT-solutions) • • Generally: Lack of imagination and bad experiences on implementing innovations	

Question 2: How could an implementation of NCM be successful in your country or organisation, and which challenges do you foresee?

Table 1	Table 2
• Case specific network building o In parallel "service blueprint" -> collaboration building o Table top session/exercises about responsibilities • Knowing your mandate (silos are important, yet blocking information transfer between disciplines)	• Norway: Success for those who embrace innovation succeed; challenge how to convince those that have not tried it, reluctance to change; change is often considered a burden • Spain: Success by bottom up + top down approach; need for approval and practitioner support; challenge: constant work due to political changes • Italy: Challenge in rigidity to change, resulting in reliance on legacy systems • Germany: Success needs not only change for management but also process for iteration; challenges in lack of digitalisation and no concept for "stay live phase"

Question 3: What further developments or advices for NCM do you suggest?

Table 1	Table 2
• Community driven design • Service blueprint with all included (all recognise) (e.g. top level C&C) • Ambassadors to show/share added value/liaison and information managers to share info • Look at UN's cluster approach • Enforce 1 system and 1 method for an entire region/country • Scientific' proof NCM works (very hard; via exercises?; ambassadors) • • Interoperability for technology & humans & procedures	• Assign roles to organisations, across all sizes of organisations • Eliminate knowledge silos o Embrace AI o Generate documents/scripts • • Gathering of relevant data o Leverage lessons learned systematically • The value of strategy-mapping

What stands out

Session 1 question 1: Do you recognise the presented issues and challenges in information and crisis management in your own professional environment? If so, how?

In response to this question the participants identified a variety of issues and challenges regarding NCM. One of the most primitive ones is the need for a system to work with. Not all countries have such a system in which they can work together with other partners during crises. Another challenge is that all parties working with the system need to be trained in the concept of NCM and the usage of the supporting system. Both the creation of systems as well as the training of professionals require a substantial amount of resources, which countries might not have or are not willing to invest. Concerns about data protection and national security as well as the fact that regulations might not be suitable yet, also form a barrier against the implementation of NCM. If laws would need to be changed, this would take up a lot of time. The general and/or professional culture can create challenges as well. Hierarchical structures present in the culture might inhibit good cooperation and professionals might be reluctant to share information. An additional challenge is how to divide the responsibilities amongst the variety of actors. Consensus is needed on who does what at which time and how to work together. This consensus can be difficult to accomplish. The degree of decentralisation is also a challenge in some countries. Decentralisation often means that a lot of small organisations need to work together. This makes it more difficult to come to an agreement with all parties to implement NCM, have well-trained professionals in all organisations involved in the crisis response and divide responsibilities between all parties.

Session 1 question 2: How do you cope with these issues and challenges?

The participants came up with a variety of coping strategies to tackle the issues and challenges they identified in the previous question. One of those coping strategies is to build trust amongst the various actors involved in the crisis response. An effective way to do so is to train and exercise together but a gentle nudge in the right direction can, at times, also do the trick. Another coping strategy is to identify common goals. When parties find out that NCM could be an asset to their as well as other parties' current operations, this can provide an incentive to start implementing it. In the case of culture being an issue for a country, the inclusion of more feminine command/leadership styles and separating command and control from information management could act as a way of coping. Participants see added value in IT-solutions and social media to improve information management.

Session 1 question 3: Would NCM fit your culture and processes? If so, how?

Most participants see the added value for an NCM-like method of information management, as quick and reliable information gathering and most important processing (preventing an information overload) is a main topic in crisis management all around Europe. Some countries have NCM-like systems and procedures in place, primarily in the Nordics. The main challenge in whether the Dutch NCM approach would fit in other countries originate in the cultural differences between the Dutch emergency services and those in other countries, and in the fact that the Netherlands has a law in place (i.e. the 2008 Safety regions act) in which shared information management processes are enforced.

Session 2 question 1: Do you recognise the presented implementation challenges?

The participants identified various challenges that come with the implementation of NCM. One of those challenges relates to culture and politics. For instance, there could be a lack of communication between organisations and/or people, a language barrier or political reasons restricting the enforcement of effective information sharing. Another challenge are the limited resources available for the implementation of NCM. A lack of collaboration between organisations and people as well as not knowing the right people, might also inhibit the implementation of NCM. Other challenges with regards to the implementation of NCM are a lack of imagination (how might this work for me in practice?) and bad experiences with past attempts to innovate or digitalise within first responder

organisations.

Session 2 question 2: How could an implementation of NCM be successful in your country or organisation, and which challenges do you foresee?

Participants mentioned that there are several ways they could see NCM become a success. One of those ways is through collaborating and exercising with other crisis response organisations, dividing responsibilities amongst them, and to keep on trying despite the challenging road ahead. Challenges that the different countries foresee in the implementation of NCM are the reliance on legacy systems, that other crisis response organisations need to be convinced to change, constant work due to political changes, a lack of or reluctance to digitalisation for crises-critical processes and a lack of trust in others in order to share information openly and speedily

Session 2 question 3: What further developments or advices for NCM do you suggest?

Several advices were gathered, which can be summarised by valuing a single system and process for all involved, in further researching the added value (and pitfalls) of increasing data-driven information management and the use of artificial intelligence, and developing network management skills (e.g. strategy mapping). Furthermore, an interest for more academic proof in the added value of NCM was recommended.

Project Symposium

The primary objective of the Symposium is to facilitate dialogue between academia and practitioners by featuring presentations and discussions on funded research projects (both ongoing and completed) and research ideas for future projects. It aims to explore promising collaboration opportunities spanning various disciplines and domains, including crisis management, assessing and mapping risks and impacts, risk and crisis communication, humanitarian logistics, AI in crisis management, socio-technical innovations, crisis management infrastructures, pandemic management, natural hazards, risks and climate change, science and policy for disaster risk management.

Each presenter delivered a 7-minute pitch in the plenary session (80 min) on, inter alia, ongoing projects, projects outcomes, overall impact and exploitation of projects' results, methodology, stakeholder involvement, existing gaps in current research, ideas for future collaboration between researchers and practitioners, etc. The pitches were followed by structured discussions (60 min). Presenters were stationed at designated tables, using a "world café style" arrangement. Audience members divided themselves among these presenter tables. This setup allowed for discussions between presenters and audience members, fostering idea exchange. The audience rotated to the next table every 10 minutes to ensure dynamic engagement throughout the session. In the following all presenters and the project abstracts (max 400 words) are listed.

BeLife: Knowledge Transfer as the Foundation of Holistic Crisis Management in the BeLIFE Project

Dr. Naomi Shulman
Technische Universität Braunschweig

The BeLIFE project, funded by the German Federal Ministry of Education and Research (BMBF), adopts a holistic approach to knowledge transfer, promoting crisis management as a continuous, proactive, and anticipatory process. BeLIFE began in October 2023 and will run for four years. It accompanies the second phase of five regional German projects, also funded by the BMBF, which are implementing solutions to prevalent issues in civil protection. These solutions respond to potential scenarios such as flooding, blackouts, and crowd panic. The five projects' planned measures seek to heighten actors' resilience and their ability to act, educate the population at large, and improve crisis planning and communication.

As the accompanying research project, BeLIFE addresses the demands of effective knowledge transfer on multiple levels and for various audiences: First, it creates an infrastructure designed to facilitate knowledge exchange between the five civil-protection projects, thus establishing an overarching knowledge space that enables each project to benefit from the others' insights into central issues and challenges. Second, it adopts the meta-perspective of sociotechnical systems to analyze how the various social and technical components of crisis management in the five projects interact and what implications this has for future crisis management, both for facilitating dialogue between researchers and practitioners and for ensuring sustainable approaches to crises. Third, BeLIFE transfers this knowledge gained from both scholarship and practice to decisionmakers and to the general population: It provides immersive insights into potential crisis scenarios at the Futures Security Lab, located at the Einstein Center Digital Future across from the German Bundestag, and disseminates key learnings from the five projects in an interactive traveling exhibit.

Ultimately, the aims of BeLIFE speak to multiple symposium topics. The project emphasizes that effective knowledge transfer, by promoting exchange between the various actors involved in crisis scenarios, is a vital aspect of the crisis management lifecycle. Moreover, it takes a holistic view of this lifecycle, demonstrating the necessary links between anticipatory planning, proactive response, and critical reflection.

The meta-perspective of sociotechnical systems as an interdisciplinary basis for knowledge transfer also prompts further avenues of inquiry. Future projects could examine how the various dimensions of knowledge transfer can help to anticipate and address challenges that arise when technical innovations encounter social needs and expectations. Such projects could focus, for instance, on how to optimize knowledge transfer between diverse actors to promote a nuanced, transparent, and responsible use of artificial intelligence in civil protection.

CREXDATA: Critical Action Planning over Extreme-Scale Data

Dr.-Ing. Jens Pottebaum
Paderborn University

Emergency management and critical action planning call for timely and accurate decision making in several, diverse applications with the goal to optimize economic, societal, or environmental impacts. Weather and health emergencies and maritime applications collect live streams of extreme scale data that reach high speeds and volumes and possess highly volatile statistical properties, incorporating uncertainties and diverse modalities through a multitude of sources. The vision of CREXDATA (https://crexdata.eu/) is to develop a generic platform for real-time critical situation management including flexible action planning and agile decision making over data of extreme scale and complexity. To achieve its vision, CREXDATA will develop a next generation Prediction-as-a-Service (PaaS) system where action planners will easily register their multimodal data stream sources and compute resource federations and graphically design predictive analytics workflows. The project will develop the algorithmic apparatus, software architectures and tools for federated predictive analytics and forecasting under uncertainty.

The purpose of the envisioned framework is to boost proactive decision making providing highly accurate and transparent short- and long-term forecasts, explainable via advanced visual analytics and accurate, real-time, augmented reality facilities. Three use cases will be used to evaluate project results: enabling authorities and first responders to act proactively in the event of weather-induced emergencies, predicting hazardous situations like potential vessel collisions at sea, and mitigating pandemic outbreaks and developing non-pharmaceutical means of treating patients. Collaborative experience will strengthen the CREXDATA project, involving 15 institutions with emergency management stakeholders in Germany, Austria and Finland. The consortium builds upon results of preceding projects like ANYWHERE, enabling impact assessments in extreme weather situation based on Multi-Hazard Early Warning System. Data from the meteorology domain like weather sensor data, nowcasts and forecasts is integrated with local data streamed from stationary and mobile sensor systems like rescue robotics.

The CREXDATA consortium will contribute knowledge along the entire chain from data processing across such extreme data up to visualization of actionable information subsuming explicit notions of uncertainty. Use case range from map-based decision support views for high level command posts to Augmented Reality tools for tactical commanders in the field. The CREXDATA architecture will incorporate ten exploitable assets based on cutting edge research, which will significantly outperform the current state of practice in respective fields. The CREXDATA project works within the greater ecosystem of HORIZON EUROPE projects that contribute to the Digital Europe Programme.

DigCBA: Digital Cash Based Assistance

Dr. Hossein Baharmand
University of Agder

DigCBA (https://digcba.uia.no/) contributes to the responsible use of digital cash-based assistance (CBA) in refugee crisis through the design, development and evaluation of evidence-based frameworks to support policy makers with the selection and use of the most suitable digital technology for delivering CBA to refugees. Digital technology, including mobile money, electronic vouchers, electronic cash, and recently blockchain-based transfers, is transforming CBA with the potential to increase access to financial resources and services during times of crisis while increasing the effectiveness and efficiency of assistance. While academic research on technology aspects of CBA is rare, the growing interest on CBA in practice has stimulated the need among CBA actors to explore using digital technology where FEASIBLE in terms of (i) contextual characteristics (e.g. data infrastructure conditions in target locations) and refugees' preference, technology access and digital literacy; and (ii) the relative readiness and acceptance of main CBA actors to adopt, implement and use the technology in target locations. The responsible use of digital CBA entails ensuring policy makers at humanitarian organizations (HOs), non-governmental organizations (NGOs), and donors make informed decisions about the use of digital technology in CBA while accounting for the contextual characteristics of the target community, and the technology readiness of involved actors. Specifically, the project takes a stakeholder-centered approach addressing the needs of beneficiaries, HOs, international and local NGOs, donors, and business partners. The project is coordinated by the University of Agder (Norway), in collaboration with NTNU, Hanken School of Economics, University Münster/ERCIS Competence Center for Crisis Management (Germany), and Makerere University Business School (Uganda). A reference group of CBA actors secures close interaction with policy makers targeted by DigCBA. The project consists of an interdisciplinary team of humanitarian operations researchers with focus on ICT, supply chain management, information systems and organization studies.

DIREKTION: Disaster Resilience Knowledge Network

Juliane Schlierkamp
Fraunhofer INT

The increasing frequency and severity of global crises, from natural disasters exacerbated by climate change to the multifaceted challenges posed by rapid technological advancements and societal shifts, necessitates a paradigm shift in disaster response and management strategies. The DIREKTION project (https://www.fire-in.eu/direktion), launched as part of the European Union's Horizon Europe initiative under the specific call HORIZON-CL3-2022-SSRI-01, aims to improve public safety and resilience against disasters. It does so by combining advanced technologies and new strategies to better manage emergencies and protect communities. DIREKTION seeks to address the critical need for an adaptive, resilient, and efficient crisis management framework that can navigate the complexities of modern disasters. It emphasizes the development of a comprehensive system that leverages the latest in research and technology while fostering collaboration across disciplines and sectors. This project underscores the potential of interdisciplinary efforts to create solutions that are not only technologically advanced but also inclusive and sustainable. This abstract outlines DIREKTION's core objectives: to harness and integrate cutting-edge technologies and methodologies to improve crisis prediction, management, and recovery processes. The project aims to innovate in areas such as real-time data analytics, communication technologies, and decision-support systems, ensuring that responses to crises are more coordinated, timely, and effective. Furthermore, DIREKTION places a significant emphasis on environmental sustainability and social responsibility, aiming to develop solutions that mitigate the impacts of disasters on communities while promoting resilience and adaptation. The methodology encompasses a blend of quantitative and qualitative research approaches, including case studies, simulations, and participatory design workshops, to ensure that the solutions developed are grounded in real-world needs and scenarios. This comprehensive approach ensures that DIREKTION's outcomes are not only theoretically sound but also practically viable and adaptable to various contexts and challenges. Anticipated outcomes of the DIREKTION project include the development of a framework for capability-driven innovation management, enhanced tools for situational awareness, and innovative strategies for community engagement and resilience building. By bridging the gap between technological innovation and societal needs, DIREKTION aims to contributing to safer, more resilient societies. In conclusion, the DIREKTION project represents a critical step forward in the quest for more effective, efficient, and equitable disaster management solutions. By aligning with Horizon Europe's vision for civil security and societal resilience, DIREKTION not only addresses the immediate challenges posed by modern crises but also lays the groundwork for future innovations in the field.

iProcureSecurity: Pre-Commercial Procurement of Innovative Triage Management Systems Strengthening Resilience and Interoperability of Emergency Medical Services

Dr. Zeynep Sofuoglu
Izmir Democracy University Medical Faculty Department of Public Health

Emergency Medical Services (EMS) in Europe are characterized by a heterogeneous landscape with diverse organizational setups, technology standards, coordination mechanisms and actors. This is the result of different historical and institutional contexts. However, these EMS are united by the common aim of providing timely care to casualties of sudden and life-threatening emergencies or disasters in cross-border settings and international humanitarian missions. Fostering the response capacities and increasing the cooperation of the Emergency Medical Services Systems is of decisive importance for strengthening the resilience of European societies.

During the prior iProcureSecurity Coordination and Support Action (CSA) project, many EMS were involved to identify, evaluate, and prioritize future challenges and needs. The creation of an interoperable, flexible triage management system supported by modern technologies was among the most requested solutions in the context of security-related scenarios.

This iProcureSecurity Pre-Commercial Procurement (PCP, https://www.iprocuresecurity.eu/) project is a result of those intense participatory processes. The action will lead to an innovative triage management system that supports the EMS on the field in the management of the Mass Casualty Incidents (MCIs). The envisaged solution will be extremely flexible and interoperable, to adapt to the different procurers' needs.

Overall, the iProcureSecurity PCP Solution aims to: (a) Fully digitalize the triage management operations on the field and provide guidance and support to the EMS practitioners; (b) Facilitate the communication and the information sharing among EMS on the field and external stakeholders. (c) Facilitate the decision-making process of the actors involved, by proving them with key information in real time, as well as with alerting and notification systems.

Following the EC Guidelines on PCP, through a competitive series of design, prototype and pilot steps, the iProcureSecurity PCP will contract suppliers to deliver the creation and deployment of the envisaged triage management system.

LINKS Strengthening links between technologies and society for European disaster resilience

Prof. Dr. Ir. Kees Boersma
Vrije University Amsterdam

Decades of research and experience have demonstrated the positive impacts of an all-of-society approach for involving members of the population in disaster risk management (DRM). In recent years, EU funding calls (e.g. H2020 and Horizon Europe) and projects for research into societal resilience, have also placed increasing emphasis on the importance of inclusive engagement with different stakeholders in DRM, including formal authorities and response organisations, businesses and private actors, and (spontaneous) volunteers and citizens, and researchers. Such broad-based participation from society requires rethinking of organisational routines and processes in crisis management, as well as innovative approaches for communication and engagement with citizens including the most vulnerable. It requires time, trust, and finding new ways to dialogue between stakeholders, both in researcher and practice.

This pitch focus on the experience and outcomes from the H2020 project LINKS (https://links-project.eu/), which concluded in 2023. The project examined the ways in which social media and crowdsourcing could be used to enhance DRM efforts in various hazardous environments. Overall, the project produced valuable lessons learned for involving and collaborating with crisis management practitioners (and other stakeholders) in research, including concrete examples of exploitation and impact of project outcomes on practitioner organisations. The pitch will also discuss how the project outcomes fed into ongoing initiatives between researchers and practitioners in projects such as the Horizon project SYNERGIES, and provide some suggestions for strengthening future research collaborations between these stakeholders.

RESCUE-MATE (1): New approaches to dynamic situational awareness and emergency responders support

Luisa Knoche
Federal Agency for Technical Relief

Storm surges are a natural hazard that is expected to increase in frequency and severity in the coming decades due to climate change. The RESCUE-MATE project (https://www.rescue-mate.de/), which is funded by the Federal Ministry of Education and Research, will address the scenario of a storm surge in Hamburg with a water level exceeding 7 meters above sea level. This could result in flooding of large parts of the port and districts close to it. All available relevant data sources must be integrated when dealing with severe storm surges due to the limited information typically available. This includes all data collected by emergency services via control centers and on site, data from sensors on site such as traffic cameras, terminal cameras, and environmental sensors, as well as all relevant information from existing systems. Additionally, information from citizens on social networks may also be considered. By collecting all of this information, it is possible to make a comprehensive assessment of the situation and derive more efficient and effective assistance. The project aims to automate, accelerate, and enrich the flow of information between authorities, emergency services, and aid organizations. Additionally, it aims to automatically collect and evaluate relevant information for assessing the situation and process it for different target groups. The project aims to develop an open, decentralized, and secure AI and data platform for situation and decision support. The platform will enable effective information exchange and generate digital twins as images of the situation on the ground, without replacing existing systems. The platform integrates public data sources and new sensor technology developed in the project, making them available to the different stakeholders.

The project addresses especially, but not only, the ISCRAM topics AI in Crisis Management and Risk and Crisis Communication. The AI-powered RESCUE-MATE system analyses available data to provide first responders with relevant, summarized information and prevent them from being overwhelmed by data. The exchange of data between organizations involved in disaster response is one of the platform's main objective.

RESCUE-MATE is set to conclude in 2027. However, further research is required to investigate short-term simulations of hazardous events and their impacts. By providing probable scenarios, emergency responder organizations can take necessary anticipatory measures, such as issuing evacuation orders and constructing flood barriers

RESCUE-MATE (2): Digital Twins and Analytical Modeling for Crisis Management: Enhancing Situation Awareness in Urban Environments

Dr. Bita Vogler
HafenCity Universität Hamburg (HCU)

As the impacts of climate change intensify, cities like Hamburg face increasing risks from natural disasters such as storm surges. In response, the RESCUE-MATE project (https://www.rescue-mate.de/), conducted within the Digital City Science (DCS) chair at HafenCity University Hamburg (HCU), integrates cutting-edge technologies, particularly Digital Twins in the Urban Space and Analytical Modeling and Simulation, to enhance crisis management strategies.

Digital Twins offer a revolutionary approach to understanding and responding to urban crises. By creating virtual replicas of physical environments, Digital Twins enable real-time monitoring, analysis, and prediction of crisis scenarios. In the context of RESCUE-MATE, Digital Twins serve as dynamic representations of vulnerable areas, infrastructure, and population density, facilitating proactive decision-making and resource allocation.

Analytical Modeling and Simulation further augment crisis management capabilities by providing insights into the potential impacts of various scenarios. Leveraging advanced algorithms and data analytics techniques, analytical models can simulate the behavior of complex systems under different conditions, allowing authorities to anticipate outcomes and plan accordingly. In RESCUE-MATE, these models enable authorities to assess the effectiveness of evacuation plans, infrastructure resilience, and resource distribution strategies, ultimately enhancing preparedness and response efforts and they can gain unprecedented insights into crisis scenarios, enabling proactive and effective responses to mitigate risks and protect lives and infrastructure.

At the conference, we propose to present the findings and methodologies developed within RESCUE-MATE Project in DSC, focusing on the role of Digital Twins and Analytical Modeling in urban crisis management. We will demonstrate how these technologies enable authorities to visualize, analyze, and simulate crisis scenarios, providing valuable insights for decision-makers and first responders and showcasing case studies and real-world examples from the RESCUE-MATE project. Additionally, we will discuss the challenges and opportunities associated with implementing these technologies, as well as future directions for research and development in the field of urban crisis management.

In summary, our conference proposal seeks to showcase the transformative potential of Digital Twins and Analytical Modeling in enhancing situation awareness and crisis response capabilities in urban environments conducted within the Digital City Science chair at HCU. We look forward to sharing our insights and collaborating with fellow researchers and practitioners in the field.

SUNRISE: Strategies and Technologies for Resilient CIs in Europe

Stefan Schauer
AIT Austrian Institute of Technology

European societies and economies depend on the reliable provision of vital services. Such vital services are delivered through Critical Infrastructures (CIs) which include basic supply networks (water, electricity), networks of information and communication services, social services (e.g., care of children or elderly), as well as complex networks of essential supply to the population (e.g., medical supply, food supply). This means that European CIs must be resilient, i.e., able to accommodate constantly changing risks and recover quickly from expected or unexpected disruptions. However, infrastructure systems and their components are interdependent. If a disruption or failure occurs in one system, it can spread to other systems and subsequently amplify the impact of a disruption many times and impact further systems.

The COVID-19 pandemic has highlighted the importance of the continuity of these vital services and has shown the need to work together for the common good. It has proven that a pandemic is not only a health crisis that disrupts CIs but underlines the extremely important link between the resilience of these CIs and our societies. The economic crisis caused by the pandemic also provides a unique opportunity to jointly 'build back better' with the focus on sustainability and green recovery.

In this context, the SUNRISE project (https://sunrise-europe.eu/) facilitates active collaboration of CIs across Europe to share best practices and jointly tackle future pandemics. Therefore, SUNRISE aims to identify pandemic-specific vital services and CIs, their dependencies, risks, cascading effects, and effective measures to tackle them at European level. This will be done by developing a comprehensive strategy for ensure the availability, security, robustness and resilience of European CIs. Further, four innovative tools are developed that include solutions for risk-based access control, resource demand prediction and management, cyber-physical resilience, and remote infrastructure inspection to improve trustworthiness, cost-effectiveness, climate-friendliness, and continuity of pandemic-specific vital services in Europe.

SUNRISE pilots the results in operational environments of the CIs while tackling some of their biggest pain points exposed by the current pandemic, promoting the approach across Europe to ensure a united front and future resilience of CIs. Overall, the SUNRISE project carefully considers legal, ethical, societal, economic, and climate aspects, ensuring that the results address not only the needs of the CIs, but also those of our society.

SUNRISE has a special focus on pandemics, but the general approaches and strategies developed in the project could also be transferred to and applied in other domains. For example, when looking at natural and technological (NaTech) disasters, it would also be necessary to integrate different points of view into one assessment and to explicitly take into account their cascading effects. With todays' interlinked and thus highly sensitive supply chains, the direct and indirect effects of such a disaster on society and social life are also an important part that needs to be considered.

Track Chairs

Abigael Collier, Medical College of Wisconsin

Amanda Lee Hughes, Brigham Young University

Amin Maghsoudi, Hanken School of Economics

Anastasios Karakostas, DRAXIS Environmental

Andrés Díaz López, Arizona State University

Andrew Arnette, Virginia Tech

Anne-Marie Barthe-Delanoë, IMT Mines Albi

Apoorva Chauhan, Concordia University of Edmonton

Aurélie Montarnal, IMT Mines Albi

Bart van Leeuwen, Owner at Netage B.V., Amsterdam-Amstelland FD

Bas Lijnse, Netherlands Defense Academy

Björn JE Johansson, Linköping University

Caroline Rizza, I3-Telecom Paris, Institut Polytechnique de Paris

Christian Elsenbast, Fraunhofer IESE

Christian Kalla, University of Münster

Christian Reuter, Technical University of Darmstadt

Christoph Lamers, Joint Competence Center for Civil Protection

Christopher Zobel, Virginia Tech

Claudia Berchtold, Fraunhofer Institute INT

Cody Buntain, University of Maryland

David Martin, Pau Costa Foundation

Diego Pajarito Grajales, University of Glasgow

Duygu Pamukcu, HEC Montreal

Filippo Sanfilippo, University of Agder

Gionata Salvietti, Università degli Studi di Siena

Hongmin Li, California State University

Hossein Baharmand, University of Agder

Jaziar Radianti, University of Agder

Jeannette Annies, University of Stuttgart

Jose J. Gonzalez, University of Agder

Josey Chacko, Mount St. Mary's University

Joy Garfield, Aston University

Julie Dugdale, University Grenoble Alps

Jürgen Moßgraber, Fraunhofer IOSB

Kayvan Yousefi Mojir, University of Borås

Kees Boersma, Vrije Universiteit Amsterdam

Laura Boyle, Queen's University Belfast

Laura Petersen, UIC Paris

Max Kinateder, NRC Canada

Melanie Reuter-Oppermann, University of Twente

Michael A. Erskine, MTSU

Michael Middelhoff, University of Münster

Muhammad Imran, Hamad Bin Khalifa University

Nicola Rupp, Deutsches Rettungsrobotik Zentrum e.V.

Niki Matinrad, Linköping University

Noureddine Bénichou, NRC Canada

Olaf Steenbergen, Netherlands Red Cross

Ophélie Morand, I3-Telecom Paris, Institut Polytechnique de Paris

Paul Geoerg, Vereinigung zur Förderung des Deutschen Brandschutzes e.V.

Peter Berggren, Linköping University KMC

Reem Abbas, Auckland University of Technology

Rob Grace, Texas Tech University

Samanta Varela Castro, Center for Research & Teaching in Economics

Shelby Knudsen, Trilateral Research

Sofie Pilemalm, Linköping University University of Agder

Steve Gwynne, Movement Strategies

Su Anson, Trilateral Research

Terje Gjøsæter, University of Agder

Weiqin Chen, Oslo Metropolitan University

Zeno Franko, Medical College of Wisconsin



Reviewers

(Completed at least 1 external review for ISCRAM 2024)

Abdoulaye Diomande, Virginia Tech

Adam Widera, University of Münster

Alexei Sharpanskykh, TU Delft

Alfredo Moreno, Universidad del Norte

Ankit Agarwalla, Virginia Polytechnic Institute and State University

Anirudh Ravi, University of Münster

Arthur Rohaert, Division of Fire Safety Engineering, Lund University, Lund, Sweden

Aura Jalal, HEC Montreal

Autumn Clark, Virginia Tech

Badreddine Farah, University of Orléans

Behnam Malmir, Georgia Southern University

Björn J.E. Johansson, Linköping University

Camilo Palacio, Austrian Red Cross

Caroline Jagtenberg, Vrije Universiteit Amsterdam

Chiara Fonio, ICONS

Enrico Ronchi, Lund University

Erik Borglund, Mid Sweden University

Firoj Alam, Qatar Computing Research Institute

Florentina Hager, University of Twente

Florian Nitschke, Lülf+

Frank Rabena, Virginia Tech

Franziska Metz, University of Twente

Georg Neubauer, Austrian Institute of Technology

George Braga, Assessoria Técnico-Administrativa do Gabinete do Comandante-Geral

Guillermo Rein, Imperial College London

Gurt Ge, University of Central Florida

Gyöngyi Kovács, Hanken University

Henrik Andersson, Norwegian University of Science and Technolog

Jason Deane, Virginia Tech

Jayr Pereira, Neural Mind

Joelle Cormier, HEC Montreal

Julie Paquette, HEC Montreal

Karen Boyce, Ulster University

Katrina Petersen, Public Safety Communication Europe

Kees Boersma, Vrije Universiteit Amsterdam

Keng-Chi Chang, University of California, San Diego

Kjetil Rustenberg, University of Agder

Kristine Steen-Tveit, University of Agder

Laurent Dubost, ATHEA

Lena-Maria Öberg, Mid Sweden University

Maedeh Sharbaf, HEC Montreal

Marcel van Berlo, TNO

Marcus Wiens, TU Bergakademie Freiberg

Maria Alejandra Rodriguez Sossa, The University of Melbourne

Marion Rauner, Universität Wien

Martyn Amos, Northumbria University

Maschio Fernando, Virginia Tech

Maxine Berthiaume, University of Ottawa

Melanie Reuter Oppermann, University of Twente

Michael King, George Mason University

Michael Kinsey, ARUP

Michelle Seref, Virginia Tech

Mike Gordon, Virginia Tech

Milad Baghersad, Florida Atlantic University

Milad Haghani, University of Sydney

Nawal Chanane, Auckland University of Technology

Nicklas Ennab Vogel, Linköpings universitet

Nikolaus Furian, TU Graz

Norun Christine Sanderson, OsloMet

Olivier Galichet, Ecole Nationale Supérieure des Officiers de Sapeurs-Pompiers

Onur Seref, Virginia Tech



Pablo Giaccaglia, Polytechnic University of Milan

Peter Berggren, Linköping University

Philipp Seeberger, Technische Hochschule Nürnberg

Pieter L. van den Berg, Rotterdam School of Management

Rizwan Sadiq, Central Asian University

Robert Larribau, Hôpitaux Universitaires de Genève

Robin Buter, University of Twente

Ruggiero Lovreglio, Massey University

Ryan Williams, Virginia Tech

Sandra Wissing, FH Münster

Soudabeh Dinani, Kansas State University

Stephen Wong, University of Alberta

Steven van Campen, Trimension

Tobias Andersson Granberg, Linkoping University

Toshihiro Osaragi, Tokyo Institute of Technology

Valérie Bélanger, HEC Montreal

Varada Krishnaswamy, Virginia Tech

Viktor Sköld Gustafsson, Linköping University

VittorioNicoletta, HEC Montreal

Wael Alkusaibati, Linköping University

Xiaojun Ruan, California State University, East Bay

Yingjie Li, Westlake University

Zeb Bowden, Virginia Tech

Zijun Long, University of Glasgow

ISCRAM 2024 Best Paper Awards

Awards Committee Chair: Filippo SanFilippo, University of Agder, Norway

Awards Committee: Jaziar Radianti, University of Agder; Rob Grace, Texas Tech University

Best CoRe Paper Award

Understanding Crisis Preparedness: Insights from Personal Values, Beliefs, Social Norms, and Personal Norms

Authors: Felix Burger and Jasmin Haunschild

Best CoRe Paper Nominees:

- Accounting for systemic risk in analysis of crisis scenarios, Jose J. Gonzalez, Reem Abbas and Sigurd Paulsen
- Enabling Information Transmission in Low-Throughput Wireless Channels for Aerial Disaster Monitoring Systems, Julian Zobel and Ralf Steinmetz
- Psychosocial Aspects of the 2021 Flood Disaster in Anonymous District, Germany: Strains, Resources, and the Role of Social Media, Francesca Müller, Marvin Kubitza and Fiedrich Frank

Best Student Paper Award

Adapting PLUM: Earthquake Early Warning with Node-Level Processing in New Zealand

Authors: Chanthujan Chandrakumar, Marion Lara Tan, Caroline Holden, Max T. Stephens and Raj Prasanna

Best Student Paper Nominees:

- Using Optimization Modelling to Determine Locations for Aerial Firefighting Resources Based on Wildfire Risk Data, Viktor Sköld Gustafsson, Tobias Andersson Granberg and Martin Waldemarsson
- Human Values for Responsible Decision-Support for Fire Services, Tina Mioch, Huib Aldewereld and Mark A. Neerincx
- Engaging the Public in Forest Fire Awareness through a Mobile Application, Emil Gatial, Zoltán Balogh, Sepideh Hassankhani Dolatabadi, Mariana Soledad Molina, Timo Kasig, Kayvan Yousefi Mojir, Marcin Wawrzyniak, Dominika Grunwald and Marcin Przybyszewski

Best WiP Paper Award

Selection of rural villages as volunteer centers for emergency response

Authors: Niki Matinrad and Tobias A. Granberg

Best WiP Paper Nominees:

- Utilizing Social Media and Crowdsourcing for Crisis Management and Communication: proposing an Integrative Framework, Kees Boersma, Nathan Clark, Robert Larruina and Emmeline Roeloffs
- SILVANUS Dashboard for Wildfire Management, Dominika Grunwald, Marcin Wawrzyniak, Marcin Przybyszewski, Emil Gatial, Zoltan Balogh and Sepideh Hassankhani-Dolatabadi
- Navigating the Digital Technology Behavior for the Distribution of Humanitarian Cash-Based Assistance: Aid Recipients Experience, Amin Maghsoudi and Abdelsalam Hamid Abakar
- Tweeting Through the Flood: Application of BERT Topic Modeling for a Comparative Flood Communication Analysis, Christin Salley, Nathan Fox and Alyssa Schubert

Best Reviewer Award

Katrina Petersen, Public Safety Communication Europe
Steven van Campen, Trimension
Keven Wesendrup, University of Münster
Jürgen Reuter, Fraunhofer IOSB

Best Reviewer Award Nominees:

- **Bas Becker**, Vrije Universiteit Amsterdam
- **Maxine Berthiaume**, University of Ottawa



Practitioner Track

Track Chairs

Nicola Rupp, Deutsches Rettungsrobotik Zentrum e.V.

Bart van Leeuwen, Owner at Netage B.V., Amsterdam-Amstelland FD

Christoph Lamers, Joint Competence Center for Civil Protection

Following the ISCRAM tradition and the conference vision, crisis management practitioners are invited to present their work in the practitioner track. Unlike other tracks, which require a CoRe- or WiP-paper for presentation, ISCRAM 2024 is adopting a novel approach informed by feedback from practitioner communities. We encourage practitioners to share their experiences, realities, and insights on current ISCRAM challenges and solutions through talks, without the need for paper submissions. This initiative aims to facilitate the sharing of pertinent ISCRAM topics and foster stronger collaboration between academics and practitioners. The submitted and accepted practitioner talk abstracts are listed in the following section.

Optimizing the deployment and use of new technologies in Search and Rescue: A recap from recent missions

Harris Georgiou
Hellenic Rescue Team of Attica (HRTA)
harris@xgeorgio.info

Alexis Vlachopoulos
Hellenic Rescue Team of Attica (HRTA)
nosailor01@gmail.com

Anastasia Andriopoulou
Hellenic Rescue Team of Attica (HRTA)
aandriop2@gmail.com

In Search and Rescue (SAR) operations, the First Responders (FR) are called upon to act in a wide range of environments, conditions and limitations, associated to both external and internal factors. Weather, location, access, infrastructure, travel time, are only some of these external factors, while team's readiness level, organization, skill level and experience, toolkits available, are some of the most important internal factors. For Urban Search and Rescue (USAR) the most critical factor, besides FR safety, is speed. Every USAR mission is an effort against time, beginning as a speed race during the first few hours after the disaster event and gradually becoming a marathon run as the hotzone gets cleared in a more focused way. In this presentation, four recent Search and Rescue (SAR) deployments of our team in 2023 are briefly presented as baseline for contextualizing the challenge of designing the next generation of assets and tools for the field operations. Namely: (a) earthquake in Turkey-Syria, (b) train crash in central Greece, (c) wildfires in Greece, (d) floods in central Greece. Lack of preparedness, coordination and poor perimeter monitoring is, as expected, a major factor in large-scale disasters. Specific differences in task prioritization and speed are evident between small-scale events like train accidents and large-scale events like devastating earthquakes, leading to corresponding changes in the deployment or even the planning phases of a SAR mission. Citizens and spontaneous volunteers are important assets often overlooked. The "last mile gap" and changes in First Aid / medical response are also crucial factors that emerge more evidently as the scale of the disaster increases. In the non-operational context, a few important guidelines and good practices can be proposed for SAR-related R&D projects. Specifically, several limitations and hindering factors in developing new technologies may work counter-productively, often very early on while in design or prototyping phases. These include very limited or no provision budget-wise for full IPR protection of the developed technologies, the low Technology Readiness Level (TRL) expectancy, limited focus to standardization and interoperability outcomes, as well as the inherently limited / highly specialized FR "market" for these products and tools. Based on the experience from DRS cluster projects of EU Horizon H2020 in recent years, important guidelines and good practices can be formulated in relation to expectation management for both FRs and technical partners, early end-user involvement in all the phase and extensive field trials throughout the projects' lifetime.

How unmanned systems empowered civilian command systems and strategic response: 15 years of lessons learnt in wildfires fight

Jean-Paul Monet
Lieutenant-colonel of Bouches-du-Rhône Fire Department (BDRFD),
Marseille, France
jean-paul.monet@orange.fr

Eric Rodriguez
Lieutenant-colonel of BDRFD, division chief
Marseille, France

Emmanuel Dombre
Emeritus CNRS research director, LIRMM, Department of Montpellier University
Montpellier, France

Marc Dumas
Lieutenant-colonel of BDRFD, Assistant chief officer for operations,
Marseille, France

For many years, military, security and safety services have been experiencing and finally using unmanned systems. The presentation's aim is to describe a short history of these robots and depict how are there employed today, in the field of civil safety, peculiarly in wildfire fight. After more than 15 years' experience of aerial imagery use, and many ground robotics trials, Bouches-du-Rhône fire department's unmanned tools team can be now share lessons learnt, in this very special context. Furthermore, authors eventually explain how this new deal implies some changes to integrate it in the incident command system processes and in decision making.

About ground robotics (UGVs) Counter Improvised Explosive Devices crews were the first to acknowledge these unusual tools. After 90ies, solutions of ruggedized, versatileand multi terrain, robots progressively emerge. Unmanned Ground Vehicles (UGVs), and this generation of tools move from C-IED missions towards CBRNe sector market, then recently in generic fire fighting and Oil and gas domains. Right now, wildfires applications revealed poor benefits, mainly due to water supply difficulties, and accessibility issues. However, some real specific cases will be reported.

About drones (UAVs)

On another hand, large employment of Unmanned Aerial Vehicles (UAVs) during Afghanistan and Middle East conflicts definitely consolidates concept and machines and allowed to develop interfaces and sensors for military uses. In the same time, in another world, leisure domain, huge number of inventors created -most frequently in their own garage! - all over the word, the large family of light drones. Nowadays, allclasses of UAVs are now fully employed, in tourism, industry, sensitive area survey, natural disaster fighting and law enforcement missions. More in our focus, Civil Protectiontroops, all over the world currently use all range of UGVs and UAVs, especially in wildfire fight. Mediterranean wildfires experiences will be reported: achievements and doubts linked to several constraints.

The conclusion will depict how the update of the command-and-control structures is paramount to increase the performance by using these tools, and how incident commanders' behavior and thinking has been influenced.

Opportunities And Challenges Raised By Introducing Robotic Systems Into The Crisis Management Process

Ivana Kruijff-Korbayová
German Rescue Robotics Center
Ivana.kruijff@rettungsrobotik.de

Firefighters and other first responders often operate in partially known and uncertain situations involving high risk due to, e.g., instability of afected structures; presence of potentially explosive or other known and unknown hazardous materials; changing situation dynamics, such as unexpected fire spread due to radiant heat transfer. This is particularly the case in medium- to large-scale incidents, such as major fires, floods or earthquakes. To reduce the danger to first responders and extend their operational capabilities, disaster response missions are increasingly supported by rescue robots, i.e. unmanned air, ground or water vehicles. Remote reconnaissance and search by robots carrying visual, thermal and other sensors enables better situation assessment and therefore better mission plannning. Remote physical operations, such as firefighting, ventilation, manipulation or heavy load transport help reduce human risk exposure. While rescue robots unarguably offer great potential benefits for mission safety and efficiency, integrating them into firstresponse teams also gives rise to various challenges. Based on practical experience from a variety of field tests and real deployments carried out with and by practioners (mostly firefighters) this talk will highlight some of the implications that the use of robots as part of incident response has on the roles, tasks and processes. In particular, (what) new roles are needed in connection to operating robots; (how) the tasks done by robots need to be reflected in mission tactics, in some cases entailing radical changes; (how) robot-teams need to be incorporated into the command structure; (what) data collected by robots needs to be transmitted and processed to derive actionable information to be presented at the right time and in the right amount to the right people; (how) data/information pull and push need to be reclected in the decision making processes. While the focus is on the (first) response phase, the need for and availabilityof data/information also needs to be taken into account in the preparedeness and recovery phases. I hope that the talk will indicate open research questions and inspire collaboration.

Engaging end users in security research

Tiina Ristmäe
THW
SYNERGISE Coordinator
Tiina.ristmaee@thw.de

Evangelos Sdongos
ASTRIAL GmbH
SYNERGISE Technical and Scientific Coordinator
e.sdongos@astrial.de

End-user involvement in security research is paramount for ensuring the effectiveness, relevance, and practical applicability of research outcomes in real-world emergency response scenarios. This abstract delves into the significance of engaging end users, such as emergency responders (ER), (from government agencies, community organizations etc.) throughout the research process in civil protection.

The culture of knowledge exchange and mutual learning has been actively practiced in the EU research project SYNERGISE. SYNERGISE develops complex technology solutions for increasing the emergency responder's safety and improving the efficiency of disaster management. In SYNERGISE project emergency responders are engaged in the research process, the scientific organisations learn about deployments and the challenges ER face during their rescue operations. To bring these different worlds together is a challenge but only the mutual learning process can bring the desired results of the research.

SYNERGISE project is user led and user oriented with defined process of user involvement. The developed technologies have to support ERs work during missions, so it is natural that ER will provide the framework of their rescue operations (scenarios, use cases, end user requirements etc.) Regular exchange between technology developers and end users is also foreseen during the prototype's development phase, in that way the connection and exchange stays alive throughout the project life time. In the validation activities at the final stage of the project end users play a crucial role – they need to evaluate the technology performance outside labs and if the systems are usable and useful during their work. Training follows these actions otherwise it wouldn't be possible to test the technologies in the field conditions.

Disaster Response by Using Location Data in Japan

Yasuhiro Soshino
Disaster Management Research Institute
Japanese Red Cross College of Nursing
y-soshino@jrcdmri.redcross.ac.jp

Masahiro Nemoto
Japanese Red Cross Hokkaido College of Nursing
nemoto@rchokkaido-cn.ac.jp

1. Background

Past disasters such as the 2011 Great East Japan Earthquake and Tsunami and the recent 2024 Noto Peninsula Earthquake revealed the challenges in identifying the passable roads, the isolated areas, and the non-officially designated shelters. To cope with these challenges, in 2018, Japanese Red Cross Kumamoto Hospital started to discuss the application of people's flow data with Agoop Corp. In the 20202 Kumamoto Floods Response, people's flow data provided by Agoop Corp. were utilized in the relief team's decision making for a rapid assessment.

2. A Field Test on the Tsunami Evacuation Monitoring

Based on this experience, the Japanese Red Cross Disaster Management Research Institute started the collaborationwiththe Agoop. Following the announcement of the damage estimation of the earthquake along withthe Chishima and Japan Trenches, the authors conducted the joint field test on the real-time tsunami evacuation during the evacuation drill in Nemuro-City, Hokkaido. In the test, the participants were asked to install a smartphone application for sharing the location with their consents while the personal data were anonymous. The collected data was analyzed to show (1) the quick identification of the evacuation sites, (2) monitoring of the evacuationstatusof the participants. The resultsof the analysiswere shown on twomonitors at the disaster response headquarters so that the local authorities and relief agencies could understand the evacuation status. In addition, after the evacuation, the participants reflected their evacuation behaviors by watching the results shown in the monitors.

3. The 2024 Noto Peninsula Earthquake Response

The recent 2024 Noto Peninsula Earthquake causedthe severe damage in infrastructure such as road, tunnels, water supply and drainage. The authors and Agoop Corp. jointly startedthe remote assessment of the disaster affected areas. By using the location data of the smartphones in the disaster affected areas, the unpassable roads were identified. In addition, the data identified a lot of isolated areas in Noto Peninsula due to the landslides. Furthermore, the people's gatherings to the non-official evacuationcentres, which localauthorities did not recognize them, were identified. The information was shared with the Japanese Red Cross and other medical relief teams, and the local authority in the disaster affected areas to support relief activities.

4. Conclusion

Application of people's flow data were effective in the rapid assessment in emergencies. For the effective and smooth application of the data, building up the information sharing system is crucial.



Command & Control Studies

Track Chairs

Björn JE Johansson, Linköping University

Peter Berggren, Linköping University KMC

WiP Paper – Command & Control Studies
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Institutional amnesia during large crisis: Effects of poor documentation

Erik A.M. Borglund
Mid Sweden University
erik.borglund@miun.se

Kjell Lagerström
Mid Sweden University
kjell.lagerstrom@gmail.com

ABSTRACT

In this "work in progress" research, we attempt to understand and explore the effects of inadequate documentation during crises. We have studied the work of three different organizations during the first six months of the Covid-19 pandemic and have been able to ascertain that almost no official documents from the organizations' crisis management teams can be found. The consequences of inadequate documentation resulted in challenges when learning from mistakes, difficulties in real-time learning, and a lack of understanding regarding the rationa le behind the decisions made. This institutional amnesia is problematic for many reasons and needs to be further studied. The next step in our research will be to study information culture and try to understand the effect of this amnesia on resilience.

Keywords

Crisis management, Covid-19, Documents, Institutional amnesia, Records.

INTRODUCTION

During large crises, command and control (C2) work is meticulously documented in various forms. These documents and records, some of which are more fragile and temporal while others are rather legally required, play a pivotal role in creating the common operational picture (COP) (Borglund, 2020). Documents and records in crisis management play a large role in creating the common operational picture (Borglund et al., 2014; Borglund, 2011). However, there are numerous instances where the absence of documents obstructs understanding of what transpired during and after a crisis, emphasizing the critical need for comprehensive documentation.

For example, when the Corona Commission, a large national assessment and review of how the Swedish government had managed Covid-19, presented their first report, much of the criticism was based on the absence of documents (Statens offentliga utredningar, 2022a, 2022b). The most severe criticism of the Swedish government was a) a lack of proper document management within the Swedish Government Office, b) difficulties accessing public records produced by the Swedish Government Office. The Corona Commission implicitly criticized the Swedish Government Office for having a very poor document and record management function during the management of the pandemic. However, there were other large crises that occurred before Covid-19 that were also criticized for poor documentation. In 2014, a large forest fire took place in Västmanland County, where 37,000 acres of forest were on fire. The fire lasted for many weeks and identified several weaknesses in the Swedish emergency management system. In the aftermath of the forest fire, one severe criticism was the absence of documentation from the authorities involved (Skogsbrandsutredningen, 2015). For example, during the forest fire, hundreds of volunteers were helping the authorities with different tasks, but in the official records there was little evidence of these activities (Sperens, 2016). In 2018, Sweden had a very dry summer and several large forest fires challenged the Swedish emergency management system, and even if authorities had improved the documentation of their activities, the lack of proper documentation was once again criticized (Statens offentliga utredningar, 2019). In the reports on the fire in 2014 and the fires in 2018 (Skogsbrandsutredningen, 2015; Statens offentliga utredningar, 2019), it is obvious that the review and assessment of how the authorities acted during the crisis could not rely on the evidential value that public records bring. The review rather relied on interviews with

people involved and their memories about what happened, instead of evidential records of the activities that took place. Noticeable is that we see evidence of activities that not are recorded at all. Lack of documentation makes it difficult to improve crisis management in the future. One area where we lack experience is understanding how information sharing worked and shared situation awareness took place in a large crisis so that challenges could be detected, and a system developed.

In this paper we investigate how authorities in the public sector fail to document their work in a command and control setting during a large crisis. The research question that guides this research is*: "What challenges arise from inadequate document management capacity during a large crisis?* This is ongoing research, and this paper presents the first findings from the work.

The paper is organized as follows. First, an overview of document management during crises is presented together with the concept of institutional amnesia. This is followed by the research method that was applied , after which the preliminary results are presented, followed by a discussion and a tentative outline for the continuation of this research.

DOCUMENT MANAGEMENT AND RECORDS MANAGEMENT DURING CRISIS

The Swedish emergency management system is complex and divided into a local, regional, and national level (Myndigheten för samhällsskydd och beredskap, 2014). There are three general principles that should guide the Swedish emergency management work:

• *The principle of responsibility means that actors retain their ordinary responsibilities in situations of crisis and disaster. This principle also includes a responsibility to support other involved parties as necessary.*

• *The principle of proximity means that crises and disasters should be managed as close as possible to those primarily concerned.*

• *The principle of similarity states that the methods and structures used in crisis and disaster management should be similar to those used under normal circumstances.*

(Krisinformation.se, 2024)

In Sweden, current recordkeeping and archival management is not separated as it is in many other countries. This is based on what constitutes the democratic openness in Sweden. The Swedish openness is based on the Swedish Constitution, the Freedom of the Press Act, which declares that all citizens have free access to public records. The Swedish Archival Act makes the Freedom of the Press Act more operational, and it states that the archive consists of all the records that are a result of a public authority's activity and business. These two legislations make it possible for everyone to access and thereby understand public authorities and their business, which is very important for Swedish democracy. Public records are evidence of activities and make it possible to understand what happened afterwards. The Swedish national legislation regulates that public business should be documented, and based on the principle of similarity, public authorities involved in crisis management should act consistently, regardless of whether there is a crisis or not.

These principles indicate that authorities should maintain consistent actions during a large crisis. According to Thomassen (2001, p. 376), records are important, *"Without records, no assessment can be made of whether individuals, private organisations and public organisations have actually carried out the actions and transactions that they had to execute, or whether the actions and transactions made meet the criteria of efficiency, legitimacy or the principles of good governance, and whether they have done things which they were not supposed to do."*

Records are evidence of the past, and during challenging situations, e.g., during a large crisis, such evidence could be extremely important as a way of producing new knowledge. In 2021, the Corona Commission, presented their first report. In the report document and records management during the Covid-19 pandemic was identified as almost non-existent.

During a large crisis, authorities usually organize into some version of a C2-setting, frequently in the form of an emergency operation center, to manage the upcoming situation.

INSTITUTIONAL AMNESIA

Other researchers have studied the effect of absent documentation in public organizations. Not only Swedish authorities have identified memory loss in organizations. Pollitt (2000) argued that organizational learning is frequently discussed, but the organizational forgetting is rarely mentioned. Many argue that the archives are the memory of the past and the archives provide organizations and culture with a lasting memory (Brothman, 2001). Therefore it is strange that the modern society with its information technology cannot provide a memory that

lasts (Pollitt, 2000). According to Pollitt (2000, p. 8), some of the underlying reasons for the institutional memory loss are the following:

"…/

1. An increasing rate of organizational re-structuring, leading to a higher proportion of relatively 'new' organizations, which rapidly lose touch with the records and personnel of their organizational 'ancestors'.

2. Rapid shifts in the media in which records are held, first from paper ('the files') to electronic media, and then from one kind of software to another. With each shift, some of the past is, for a variety of reasons, likely to be lost.

3. The decline of the concept of public service as a permanent career, leading to a higher proportion of public servants with only transient experience of their organizations.

4. The popularity of ideas of unceasing, radical change. These include some of the most fashionable improvement techniques in the public sector, such as re-engineering and benchmarking. These systems of ideas include suggestions that the past is no longer relevant, or even that to look back is defeatist and dangerous.

/…"

In the work by Borglund (2020) the management of records during crisis very much relied on the discretion of individuals. Some records were necessary to keep for legal reasons, but a rather large amount of information created during large crises were never documented and kept for future use.

T he impact of memory, relevant to crisis management, can be summarized as follows, according to Pollitt (2000, pp. 13-14):

1. The learning from mistakes is rare, and organizations tend to re-try things that failed 5–10 years ago.
2. Using fashionable solutions increases vulnerability.
3. Using other organizations' solutions to move away from certain trajectories but without knowing if the solution works.
4. Lack in capacity to understand the knowledge value amongst "long-serving staff".
5. Inconsistencies and discontinuities in dealing with citizens/service users. From a crisis management perspective 1–4 are very relevant and therefore institutional amnesia should be avoided.

METHOD

Our research uses a qualitative research approach (Myers, 2009). The research can be a seen as a case study (Lee, 1989) where we have aimed to understand how the Covid-19 pandemic was managed by one municipality, one region and one county administrative board during the first six months of the pandemic. Additionally, we sought to understand and further scrutinize document management and the document management role specific to the management of the Covid-19 pandemic in one region in Sweden. The data collection has primarily focused on studying the public records registered as created or received by these three authorities. We decided to collect records related to C2 organization in the three studied organizations from when WHO declared Covid-19 a pandemic on March 11, 2020, and the six months following. The selection of these three authorities was defined as expedient selection (Hartman, 1998), where selection was meant to be able contribute to the answer of the research question. The region is a responsible for many tasks, but primarily for medical service within the region. The country administrative board is a government authority in a region that serves as link between people, municipalities, the government and central authorities. Furthermore, the county administrative board is responsible for crisis management collaboration within the region 's borders, while the municipality is responsible for crisis management collaboration within the municipality 's borders.

The three organizations studied had all activated their crisis organization, i.e., they organized according to the recommendations from MSB (The Swedish Civil Contingencies Agency).

We conducted a search for data within the registers of public records generated by the crisis organization at the regional level, the county administrative board, and a specific municipality. In each case, we focused on posts, records, or files that were tagged with metadata labels such as "crisis", "operation picture", or "crisis organization". We searched the public records registers to find the original records. We also conducted further investigations within the three organizations to understand if other records and documents could be received.

INITIAL FINDINGS

We started by searching the official registers of the three authorities. As all documents made and received from a public authority's business should be part of the authority's archives, finding the documents produced during the crisis management organisation in the authority's official register should be possible.

Municipality

In the official register of the municipality, there was only one record related to the municipality's crisis management during Covid-19, primarily related to the COP. We also used other search strings, but it did not result in any additionalofficial documents. We then asked the municipality for a manual search, and they provided six different closed records. One record could consist of one or several documents. The municipality replied with six records labelled the crisis management board (*Krisledningsnämnd* in Swedish). The reply from the head of security of the municipality did not provide us with any physical documents to study, only a list of six records. What documents these records contained was not clear. The municipality studied has a catalogue of all processes and documents that can be incorporated into these processes. The catalogue regulates how to deal with the documents, including preservation time, etc. All risk and crisis management documents in the studied municipality should be preserved for three years. Therefore, it is strange that for the first six months of Covid-19, there are only six records.

County Administrative Board

The county administrative board is an important authority in the Swedish crisis management system. There are twenty-one country administrative boards, and their tasks during larger crisis are:

- Collaborate within the county and with neighboring counties.
- Coordinate activities and information among municipalities, regions, and authorities during a crisis.
- Direct and prioritize state and international reinforcement resources based on government decisions.
- Implement preventive measures to prevent societal disruptions.

A search in the official register of the county administrative board during the first six month of Covid-19 resulted in six records with some relevance to the crisis management organizations.

Region

In the region, during a crisis, an emergency management board is activated with the task of governing the region during a ciris. As with the municipality, we searched for documents in the official register. The result was a total of 5 records containing "common operational picture", out of which some were irreleva nt. It was not possible to further investigate the relevance and the presence of documents within those records in the official register.

In a consultant report from KPMG (KPMG, 2021) the work of the entire region during the first month of Covid-19 was evaluated. In the report it is obvious that the region with its subdivisions had regular meetings related to Covid-19. Some parts of the region organized into a staff organization. The KPMG report also indicated that through there was a lack of documentation of decisions, KPMG were able to analyze more than 400 nonregistered documents such as minutes, decision logs, and memorandums.

Secret or Not Secret Documents

One possible explanation for the absence of documents from the first six month of the Covid-19 pandemic is that the documents that do exist have been secret. However, if they were secret, they should be listed in the registers. In Swedish legislation it is stated that records with a secret content must be registered, where you should find the date, the register number, who created the record, and what it is about (Larsson, 2019). This means that if the three authorities classified their records from the crisis organization and their common operation picture as secret, it should still be possible to find the records when searched for.

A NATIONAL DOCUMENTATION TASK

In connection with the first wave of the coronavirus pandemic, the National Archives urged government agencies to preserve documents containing information about the pandemic (Riksarkivet, 2020). The following points were

highlighted:

• Improving the preservation of documents such as internal communication, including emails, meeting notes, and information for employees.

• Ensuring that relevant information from government websites and social media is preserved.

In 2023, the National Archives issued a general recommendation encouraging government agencies and other entities to expand the preservation of societal information during crises. It is worth noting that the National Archives considers crises as disruptions similar to the Swedish Civil Contingencies Agency (See in page 5-6 Riksarkivet, 2023).

DISCUSSION OF THE COLLECTED DATA

While it is too early to draw any definitive conclusions, it is obvious that the criticism raised after the forest fires in 2014 and 2018 was not only relevant for the management of forest fire crisis. During the Covid -19 pandemic, thousands of people died in Sweden, which makes it clear that it was a large crisis. Why is it so difficult to find documents and records from a large crisis? At this stage we have only scratched the surface, but based on previous research (Borglund, 2017; Borglund, 2020; Holmström, 2017), the way documents are created is a possible explanation, and the lack of adequate system support to manage the created documents over time is another. Whiteboards are important in the documentation of common operation pictures, but a whiteboard and large flip chart sheets are also difficult to preserve. In the temporal location of the crisis management, the systems available are not designed to capture the records and documents created. However, it is notable that in a nother study where an organization did not establish any temporal crisis organization (like an emergency operation center) due to the need to work from home, they found that the documentation was greatly improved by working as normal (Borglund et al., 2021). Could the explanation in poor documentation be found in how the work is done?

What Challenges Arise From Inadequate Document Management Capacity During a Large Crisis?

Naturally, many challenges arise when very few documents from C2 work is found. Argyris and Schön (1974) presented Espoused Theory and a Theory-in-Use. Espoused Theory is about the theories a person believes he or she follows. But we apply the concept from (Argyris & Schön, 1974) as a concept where espoused theory is also about the action or activities a person believes he or she has done. Theory in use, however, refers to the theories a person actually uses or follows. The lack of documents from a crisis makes it almost impossible for external or internal auditors to assess how the organizations actually worked. The story, or the narrative, from the crisis, comes from the memories of the people involved. The evaluation of the forest fire in Sweden 2014, 2018 (Myndigheten för samhällsskydd och beredskap, 2016; Skogsbrandsutredningen, 2015; Statens offentliga utredningar, 2019) and the evaluation of Covid-19 in Sweden (Statens offentliga utredningar, 2022a, 2022b) a **l** indicate that the narrative of the actions comes from the memories of the people involved. I.e. the actual action and activities (theory in use) cannot be evaluated.

Another problem caused by inadequate document management during crisis is that it becomes a challenge to learn. In the double loop learning presented by Argyris and Schön (1978), documents can be important in the analysis of the results that are the effect of certain actions and assumptions. The double loop learning is visualized in fig 1.

During the winter of 2024, a large traffic incident occurred on E22 in Skåne where 1,000 vehicles were stuck on the highway due to snow for up to 50 hours. In a recently published report from this incident a lack of documentation was identified also in this case, which made it impossible for authorities to understand the situation and to share the situation awareness with other authorities. This in turn resulted in the absence of a common operational picture which caused a delay for all decisions needed (Capgemeni & Länsstyrelsen Skåne, 2024).

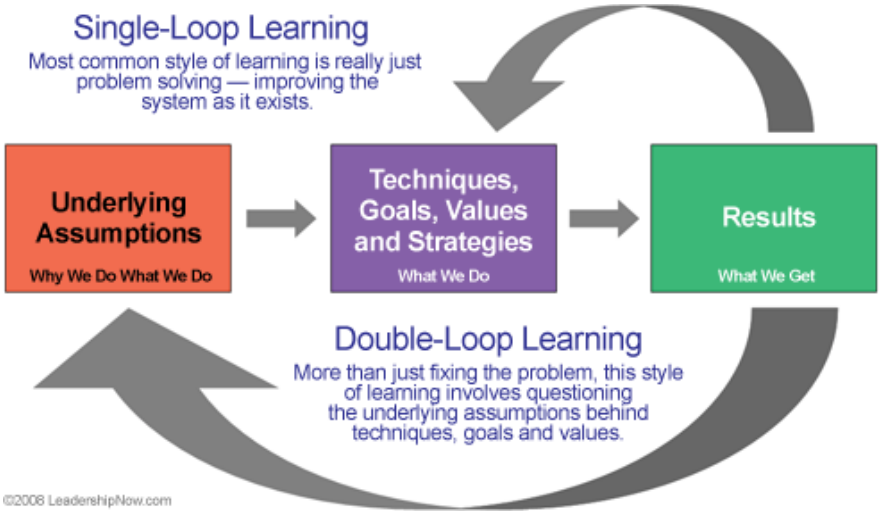


Figure 1. Double loop learning.

Category of Crisis and Its Temporality

Not all crises are the same, and an important difference between them is the need to manage the core of the crisis. In a forest fire, the problem is the fire itself, the crisis is a result of the fire. During the Covid-19 pandemic the problem was the virus. These rather obvious differences in crisis categories also affect the time frame in which the crisis management is taking place. The mentioned forest fires in Sweden in 2014, and 2018 required crisis management for a period of weeks, whereas the Covid-19 pandemic required support for years. This difference in temporality is bound to affect the documentation practice during a crisis. None of the actors involved in these crises realized that the absence of documents was an issue, nor that a poor document management should not be accepted.(Myndigheten för samhällsskydd och beredskap, 2016; Skogsbrandsutredningen, 2015; Statens offentliga utredningar, 2019, 2022a, 2022b).

CONCLUDING REMARKS

One of the most significant issues stemming from the absence of documents and records is the inability to learn from past mistakes, thereby minimizing the risk of repeating unsuccessful solutions. However, documents a lso play a crucial role in understanding the common operational picture at a given time, and the shared understanding of the situation that formed the basis for specific decisions. During command and control work , information sharing is crucial to establish a shared understanding of the ongoing situation (Sonnenwald, 2006). Effective information sharing can also result in a shared situation awareness (Chubala et al., 2023). To be able to improve the information sharing and work with a common operational picture, documents can play an important role.

There can be many explanations for the lack of documentation from the crisis management in the three studied organizations. One explanation is lack of time, and stressful work situations for the staff in the C2 setting. However, when examining the principle of similarity, it suggests that the methods and structures employed in crisis and disaster management should closely resemble those used in normal circumstances. In normal circumstances activities in public authorities are documented and registered. We believe that many of the activities during the first six months of the pandemic in the studied authorities was documented but never registered.

The next phase of this research involves delving deeper into the information culture within organizations to gain a better understanding of why documents and records are not being kept, organized, and made readily accessible. Additionally, we will assess the quality of existing evaluations of large crises. However, can we truly draw meaningful conclusions without sufficient documentation and records?

REFERENCES

Argyris, C., & Schön, D. A. (1978). *Organizational learning : a theory of action perspective*. Addison-Wesley.

Argyris, C., & Schön, D. A. (1974). *Theory in practice : increasing professional effectiveness* (1st ed.). Jossey-Bass Publishers.

Borglund, E. (2017). *The Role of Artefacts in Creating a Common Operational Picture During Large Crises* the 14t International Conference on Information Systems for Crisis Response and Management, Albi, France, May 2017.

Borglund, E., Granholm, M., & Anderson, U. (2021). *Virtual Emergency Operation Centre: How to manage a crisis from an EOC when you need to work from home* 18th ISCRAM Conference Blacksburg, VA, USA May 2021.

Borglund, E., Landgren, J., & Lintzen, M. (2014). *Lägesbilder: Att skapa och analysera lägesbilder vid samhällsstörning [Common operational pictures: To create and analyze common operational pictures during large crisis]*. Myndigheten för samhällsskydd och beredskap.

Borglund, E. A. M. (2011). Records Role in Creating a Common Cperational Picture. In Judith Molka-Danielsen, H. W. Nicolajsen, & J. S. Persson (Eds.), *Selected Papers of the Information Systems Research Seminar in Scandinavia Nr. 1 (2010): IRIS 33 Engaged Scandinavian Research* . Tapir Academic Press

Borglund, E. A. M. (2020). Records-Making During Crisis Management – Rule Based or Discretion Driven? In A. Sundqvist, G. Berget, J. Nolin, & K. I. Skjerdingstad, *Sustainable Digital Communities* International Conference on Information, iConference 2020, Borås, Sweden.

Brothman, B. (2001). The Past That Archives Keep: Memory, History, and the Preservation of Archival Records. *Archivaria*, *51*, 48-80. https://archivaria.ca/index.php/archivaria/article/view/12794

Capgemeni & Länsstyrelsen Skåne. (2024). *Oberoende Utredning av Trafikstoppet på E22, 3-5 januari 2024, Händelseutredning 2024-03-25 [Independent Investigation of the Traffic Stop on the E22, 3 -5 January 2024]*. https://www.lansstyrelsen.se/skane/om-oss/nyheter-och-press/nyheter---skane/2024-03-26-oberoende-utvardering-klar-om-trafikstoppet-pa-e22.html

Chubala, C. M., Hunter, A., Dithurbide, L., & Neyedli, H. F. (2023). Building Situation Awareness and Team Cohesion through Effective Information Sharing in a Distributed Team -Based Command and Control Scenario. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* , *67*(1), 710-715. https://doi.org/10.1177/21695067231192866

Hartman, J. (1998). *Vetenskapligt tänkande : från kunskapsteori till metodteori*. Studentlitteratur.

Holmström, L. (2017). *Dokumentationens gråzoner : Utmaningar när krishanteringen får stöd av frivilliga* [Student thesis, DiVA. http://urn.kb.se/resolve?urn=urn:nbn:se:miun:diva -31257

KPMG. (2021). *Utvärdering av covid-19 för Region Uppsala*. Region Uppsala Retrieved from https://regionuppsala.se/contentassets/cb1da4750f9c44f1858edf32f852e6a1/slutrapport_covid19_utvardering_region_uppsala_210503.pdf

Krisinformation.se. (2024). *Krishanteringens grunder [the basics of crisis management]*. Krisinformation. Retrieved 29 mars from https://www.krisinformation.se/detta -gor-samhallet/mer-om-sveriges-krishanteringssystem/krishanteringens-grunder

Larsson, E. (2019). *Måste jag diarieföra det här? : en handbok om regler och rutiner i ärenderegistrering* . SKL Kommentus AB.

Myndigheten för samhällsskydd och beredskap. (2016). *Ansvar, samverkan, handling : åtgärder för stärkt krisberedskap utifrån erfarenheterna från skogsbranden i Västmanland 2014 [Responsibility, collaboration, action: measures for enhanced crisis preparedness based on the experience of the forest fire in Västmanland 2014]*. Myndigheten för samhällsskydd och beredskap.

Pollitt, C. (2000). Institutional Amnesia: A Paradox of the 'Information Age'? *Prometheus*, *18*(1), 5-16. https://doi.org/10.1080/08109020050000627

Riksarkivet. (2020). *Riksarkivets rekommendationer till statliga myndigheter med anledning av Corona-pandemin*. Retrieved 2024-01-05 from https://riksarkivet.se/dokumentationshantering-corona-pandemin

Riksarkivet. (2023). *Utökat bevarande vid kris: Vägledning* Retrieved from https://riksarkivet.se/Media/pdf-filer/doi-t/vagledning_utokat_bevarande_vid_kris.pdf

Skogsbrandsutredningen. (2015). *Rapport från Skogsbrandsutredningen. [Report from the Forest Fire Investigation]*. Regeringen [Swedish Government]

Sonnenwald, D. H. (2006). Challenges in Sharing Information Effectively: Examples from Command and Control. *Information Research: an International Electronic Journal* , *11*(4), n4.

Statens offentliga utredningar. (2019). *Skogsbränderna sommaren 2018:Betänkande av 2018 års*

skogsbrandsutredning SOU 2019:7. Stockholm Statens offentliga utredningar

Statens offentliga utredningar. (2022a). *Sverige under pandemin : Förutsättningar, vägval och utvärdering Volym 2*. Regeringskansliet.

Statens offentliga utredningar. (2022b). *Sverige under pandemin : Samhällets, företagens och enskildas ekonomi Volym 1*. Regeringskansliet.

Civil Protection in Europe: Towards a Unified Command System? *Update on knowledge-sharing work and future actions*

Jean-Paul Monet
Senior Officer (ret.)
Bouches-du-Rhône Fire Dpt, France
jean-paul.monet@orange.fr

Anaïs Saint-Jonsson
Assistant Professor, IMPGT
Aix-Marseille University, France

Sergio Giacomo Pirone
Operation Commander, Regional Forest Fire Fighters Corp - Piedmont, Italy

Marc Castellnou Ribau
Chief of Forest Fire Division
Catalonia Fire & Rescue Service, Spain

Stéphane Poyau
Division Chief Logistics
Landes Fire Dpt, France

Marc Dumas
Head of Operational Subdivision
Bouches-du-Rhône Fire Dpt, France

Miguel Cruz
Deputy National Commander
Emergency and Civil Protection National Directorate, Portugal

Tony Johansson
Incident Commander/
Fire Protection Engineer
Södertörn Fire & Rescue Service, Sweden

Christoph Lamers
Representative of North Rhine Westphalia st.
Joint Competence Centre for Civil Protection, Germany

Christos Lampris
General Officer (ret.)
National Fire Service, Greece

Schaller Pierre
Senior Officer (ret.)
National Fire Service, France

ABSTRACT

In the context of climate change, increasing disaster severity and extension, the paper suggests to DGECHO, agency managing the European civil protection mechanism, an improvement for joint operations command-and-control practices. It constitutes an update of Monet et al. (2020) paper, enhancing the work progresses in terms of analysis, applicability, monitoring and training.

In this perspective, the work of the authors is completely dedicated to the improvement of interoperability when different (national) civil protection forces work together, to respond to emergencies or disasters. Building on

theoretical comparative studies, this opinion paper now relies on an interoperability monitoring tool and consolidates the ways on defining the future for an "European Command System", proposing innovative operational practices by associating knowledge exchange and operational training.

Actually, the presentation will depict how to organize a training strategy at European level, alternating international and regional courses, to pave the way for future harmonization.

Keywords

Crisis Management; Civil Protection; Incident; Command and Control System; Interoperability; European Union

INTRODUCTION

The aim of this paper is to propose elements of a "new governance in operations" for European civil protection. To avoid any misunderstanding on the term governance, the authors specify that the present work does not tackle the whole domain of civil protection policies. This presentation is focused on a small part of these public policies, dealing with disaster response management and associated training.

All the years, from 2016 onwards, gave many occasions to activate at large scale the EU Civil Protection Mechanism (EUCPM): floods, hurricane, earthquakes, and wildfires. Climate change and the increasing occurrence of natural risks reveal a need and the opportunity to develop a European command and control framework (Maestracci, 2013). For more than 15 years, authors, who are in a predominant majority responders, have been working in this field of international mutual aid for civil protection. This paper's purpose is to provide inputs for the further empowerment of the EUCPM.

The area of large incidents and disaster management is frequently evoked by academics (Topper and Lagadec, 2013; Dufès and Ratinaud, 2014, Wybo and Latiers, 2003), but slightly targeting incident management practices. Many collaborative EU funded projects (FP7, H2020, and Horizon Europe) related to disaster research topics are contributory too, but totally focused on innovation, mainly technique one. Even if some research exists, "real" field incident management topic remains in Europe a kind of grey zone.

Today more than ever, in a changing risk landscape, and after many destabilizing incidents (terrorism in France and Belgium; wildfires in Portugal, Spain, France, Sweden, and Greece; flood and storms), empowerment of the European civil protection mechanism is essential. And clearly under the leadership of the Directorate-General for European Civil Protection and Humanitarian Aid of the European Commission (DGECHO).

Under pressure of climate change and wildfires geographical area extension, civil protection forces and fire departments have, right away, to respond to the risk of hyper violent fire behavior (e.g. extreme wildfires events: Tedim, Xanthopoulos et al., 2018), and natural disasters due to extreme meteorological conditions. While incident management is sometimes a sovereignty issue, or perceived as it, authors suggestion is conceived as an intermediary stage towards harmonization. Hence, in this transitional period, effort is dedicated to the crucial interoperability increase. For this purpose, national systems are preserved, and suggested to be customized in a kind of "international hybridization", to be better compliant when the European civil protection mechanism is activated.

The following paper presents ongoing works, written by knowledgeable practitioners of civil protection, aiming to provide insights and opinions, to support dealing with current global trends and the reality of "terrain response". As very challenging events (Maureen Essen et al. 2022), wildfires, due to their extension and intensity, could play in this evolution, the peculiar role they already performed in some fire-prone countries (USA, Australia, Portugal, France…) in fostering command and control enhancements, further extended to all hazard management, as Rowley and others authors explain for American incident command system (ICS).

The document is organized in five parts: built on lessons learnt study and real EUCPM activations, it presents an interoperability monitoring tool supporting the evaluation of the cooperation. After a quick presentation of existing systems, the work identifies by comparison the possible contributions that each could make to new European operating modes. And eventually provides some pathways to share the knowledge and organizing training and exercising in the EU.

Finally, in a nutshell, this transitional proposal, expecting a leveraging effect, opens the way for the EU towards establishing a sui generis harmonized operational command and control system (ECS), which would be interoperable with the other major pre-existing global ones, without being enslaved to them.

TERMINOLOGY

Definitions

To clarify the paper vocabulary, we must first reach an agreement on some basic terminology, which unfortunately is not yet harmonized across the world. After discussing terminology, we will present our ideas on proceedings.

We could start with the terms "crisis" and "crisis management" themselves: these words, initially meaning situations in which decision bodies lose control and their ability to act, have over time been stripped of their original meanings and now are more generally used to mean big incidents and incident management. Big incidents are not necessarily crises in the original sense, and a misbalance between available and needed means is not enough to create an actual crisis. However, random or quick evolution, lack of decisions, emergence of power struggles, and other signals of weakness need to be tackled in order to avoid crisis occurrence. In order to characterize these issues, and because no classification system exists in Europe, the US classification of incidents (see below Table 1.) could be used or slightly adapted to provide harmonization at an international level.

Table 1 : US Incident classification

	"Definition"	**Example**
Type I	Huge incident	i.e. large on shore pollution lasting weeks P>500
Type II	Big incident	i.e. Bushfire lasting 3 days P<500
Type III	Important incident	i.e. Railway crash lasting one day with multiple assets
Type IV	First level of incident implying transfer of command	i.e. structure fire
Type V	Small usual mission dealt by one or two assets.	i.e. Emergency Medical Service mission with one ambulance

What shall we call LEMA? LEMA is the European acronym for Local Emergency Management Authority. Nowadays it is frequently used in a broad sense, meaning all local incident managing activities. In our proposal, we use the term LEMA in its original sense, to refer to an elected (or non-elected) single person authority, who is a policy maker in charge of the strategy above the Incident Commander. To draw comparisons, UK uses "Competent Authority" and US ICS "Authority Having Jurisdiction".

What is an ICP? Incident Command Post (ICP) or Field Command Post (FCP), interchangeably, is remarkably close to US terminology. This term is our favourite to describe "site headquarters' ' preferred to **OSOCC which** was initially set by UNOCHA, in UNDAC/INSARAG guidelines. Meaning On-Site Operations Coordination Centre (UNOCHA Field Handbook, 2018), this wording has been surprisingly corrupted by the term "virtual OSOCC" (1998), which describes a shared assessment IT tool. Introducing *de facto c*onfusion which might be a permanent source of misunderstanding, we suggest not using this term, and favour ICP instead.

What about the so-called Standard Operating Procedures (SOPs)? While this acronym is trendy in Europe, with different understandings, (in total absence of an harmonization framework) we propose to clarify it by introducing a distinction. We suggest using STPs for *Standard Technical Procedures* and SCCPs for *Standard Command and Control Procedures*; both constituting SOPs (cf. Figure 7.)

STPs: as all EUCPM participant states have a great reservoir of skills relating to civil protection techniques, there is an opportunity to share and exchange. Currently technical practices are not regulated by any EU or other standard, so it's now impossible to have standardized training on these practices.

SCCPs: As it has been said, there is a real need to empower EUCPM through a "European Command System" (ECS), giving harmonized command and control methods at the international level in the EU. All these practices are gathered under SCCPs acronym.

Eventually, this paper has included a Glossary for all abbreviation used, in the last part of the paper

CONTEXT, LESSONS LEARNT AND PROBLEMATIC

Large and violent wildfires in Greece (2007, 2018, 2021, 2023), Portugal (2016, 2017), Spain (2012, 2015, 2019) France (2003, 2016, 2019, 2022), Sweden (2018), Australia (2019), South Africa, Canada (2023) and California (2018), gave in the last decade several opportunities to witness climate change consequences. These events also demonstrated a real need for a larger commitment of EU practitioners, to increase the quality of European civil protection mutual support. And the recent cases of EUCP mechanism activation (2018, 2021, 2022, 2023) including floods in Belgium and Germany (2021), and France (end of 2023 to 2024) demonstrated the increasing need for field experts and very experienced responders.

The Swedish 2018 fires gave another perspective on emergency management in a "non-fire prone country", using NATO guidance to deal with the unusual situation. Actually, the Kårböle-Ljusdal fire Incident command post followed the NATO standard organization (cf. below: "NATO framework"), with a large Incident Management Team, driven by an Incident Commander, responsible for decisions, ground tactics and management. The latter was permanently reporting to the local policy maker (Province Governor), in charge of strategy, and in contact with national authorities. In this case, despite the larger EUCPM activation of its history (20 years), and the involvement of eight countries in response to support Sweden, DG ECHO constituted only a very agile structure of coordination in the host country at national HQ. And that was valuable, a kind of "proof of concept", which allowed all kinds of wildfires modules sent, to be adequately involved in the fight, and additionally, the Swedish C2S demonstrated its adaptability in efficiently integrating the foreigner experts.

Greek fires in 2021 and 2023, demonstrated a real integration of European forces by pre-positioning of some other European resources and the repartition of large and diverse involvement of EU wildfires modules and assets.

Facing all these cases, we must admit that there is a great reservoir of data in the study of these cases. Some lessons learnt and seasonal reports have been written, notably by European Commission bodies (San-Miguel-Ayanz et al., 2023), but rarely dealing with command-and-control issues.

Regardless of the cases, the **same problematic arose**: if at ground level all foreigner strike teams and platoons are integrated with relevancy by "missional assignments" in the fight to the disaster, it remains the question about the interoperability for the officers and experts, that can neither understand clearly the host nation command and control practices, nor been properly integrated in the incident management team, to provide a support at the level of their expertise.

INTEROPERABILITY MONITORING TOOL

Principle

Authors propose and use a matrixial interoperability evaluation tool which offers a simple and user-friendly framework for evaluating and enhancing teamwork to prepare for and respond to joint civil protection operations.

The tool can be used to document in a structured manner the lessons learned in the aftermath of disaster response. Alternatively, it can be used at any time to monitor existing cooperative activities (e.g., cross-border operations) to check and identify areas for improvement, building a kind of "pathway toward interoperability".

Inspired by an US[1] radio communication monitoring tool (SAFECOM program), this interoperability evaluation tool (Monet, Mannaioni et al., 2018) was initially developed, by adding others criteria, at national level to meets satisfying two aims:

- Address all aspects of national interagency civil protection operations: policy and guidance, training and communication, and command and control practices.
- Assess and score each criterium and create a combined score using weights reflecting the decided relative importance of each.

The structure of this tool was slightly modified to reflect international operations and used to assess interoperability in a 2018 European Civil Protection Mechanism (EUCPM) wildfire operation in Sweden involving both Swedish and French responders.

This work has been finalized under the auspices of a EU funded project (Monet et al. 2023)

[1] Department of Homeland Security, 2004

Depiction

The structure of the matrixial interoperability evaluation tool, dedicated for assessing a joint operation which relies on the use of the table above to assign an Evaluation Level of 1 – 5 for each interoperability component, defining 5 as the best ranking.

- Governance
- Standardized Operational Procedures (SOPs)
- Data Technologies
- Voice Technologies
- Training and Exercises
- Language
- Operations

A Component Score is calculated for each component by multiplying its Evaluation Level by a Component Weight. An Overall Score is then calculated by adding up the seven interoperability components.

It was judged that the first four components, Governance, SOPs, Data and Voice Technologies, are considered of equal importance. Training and Exercises as well as Language skills are viewed deemed as being twice as important as the first four components. Finally, the Operations component was judged to be four times as important as the first four components.

All these Component weights ensure a maximum overall Score of 100. The Overall Score is supposed to rank the performance of the cooperation of civil protection forces at incident management level.

To give a score on the efficiency of interoperability, policy makers or responders can use this innovative and simple tool themselves to improve their joint operation, identifying the way(s) of improvement. The Table 2 below presents the tool.

Table 2. Evaluation components and levels

Interoperability component	Evaluation level				
	1	**2**	**3**	**4**	**5**
Governance	National agencies working separately	No formal joint work between national agencies	Occasional bilateral or multilateral formal group for join work and missions	Multilateral agreement (s) for join work and missions	Specific bilateral agreement (s) for join work and missions
SOPs	Harmonized national SOPs	National SOPs to receive international aid	Regional SOPs	Multilateral SOPs	Specific bilateral SOPs
Data Technologies	Manual files swap	Bilateral files exchange via commercial software	Sharing by a dedicated ad hoc interface App	One-way sharing with standards-based software	Two-way sharing
Voice Technologies	Swap radio	Gateway	Shared channels	Proprietary channels system	Bi or multilateral standards-based sharing systems
Training	National unified training doctrine	National full-scale exercises	Cross bordered training and exercise	Regional training and exercising	Multilateral training and exercising
Language Skills	Use of interpreters and/or immediate translation tools (e.g., I.A., smartphones)	Basic use of third language (A)	Independent use of third language (B)	Proficient use of third language (C)	Use of host country language
Operations	Teams working separately with limited communication	Teams working together with formalized (briefing, radio) communications	Teams achieved being integrated in the local command system	International teams taking roles in local ICS	Same incident command system

Summary

These last four years, the tool was applied by authors to some real cases of EUCPM activation, and after Modex (Modular exercises, EUCPM field trainings). Even if it highlighted several criteria's room for improvement, one of them, significantly always appearing with a low score, was C2S practices.

STATE OF THE ART IN COMMAND-AND-CONTROL PRACTICES

US ICS system

In the 1970's, the Incident Command System (**ICS)** was developed by the US Forest Service to manage forest fires and other natural hazards, and it was quickly recognized as an efficient inter-agency control system. After the terrorist attacks of 11 September, following an *ad hoc* decision of the federal government, the concept rapidly and widely spread throughout the United States. Now largely used there in its "All Hazards Management" version, ICS today involves thousands of civil servants, foresters, police officers, fire fighters and elected officials who constitute more than 15 regional teams. At the highest level, responding to the largest incidents known as Type I, these "Incident Management Teams" (see IMT, Figure 1.), are constituted with 50 specialists on 24/7 duty.

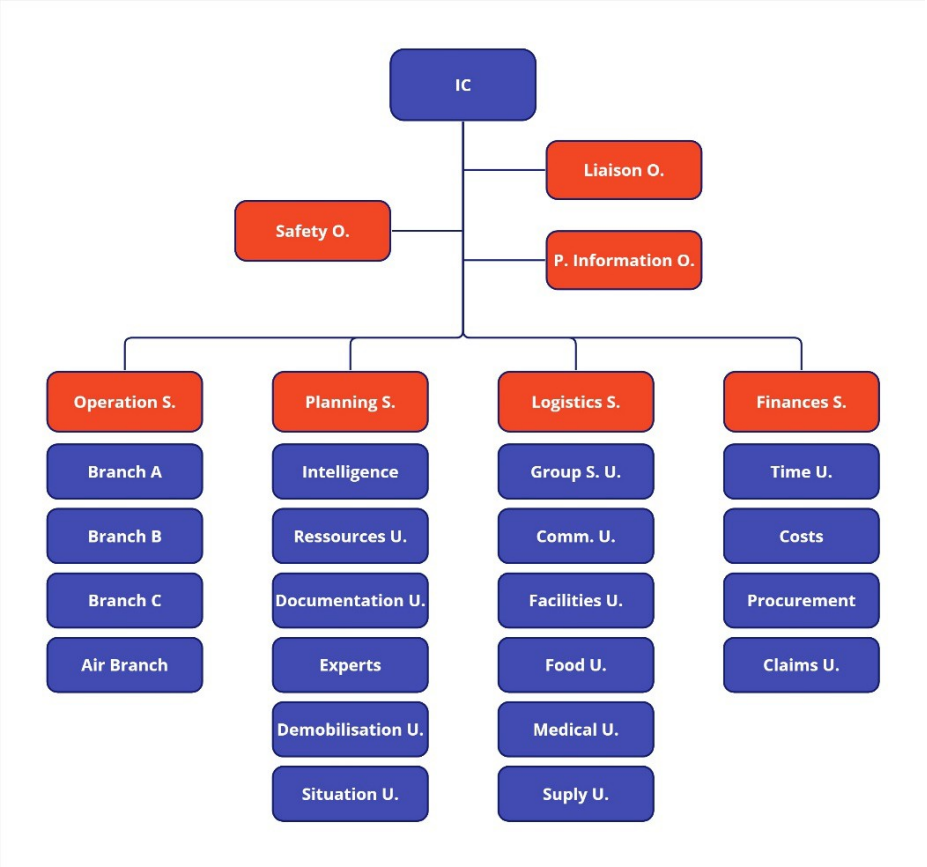


Figure 1. IMT

Today ICS is disseminated all over the world, with some adaptations. However, in France less than 20 fire officers are acquainted with this system. After being trained through the IS-420 Incident Management Team course (an intermediary rank in the ICS training program which starts from IS-100 to IS-800), a few of them participated in exercises in California and Australia. Some Spanish provincial foresters and Norwegian fire fighters are now implementing the US ICS directly. The fire service of Catalonia (GRAF, Grup d'Actuacions Forestal) is also influenced by ICS, especially in the practice of use of fire, which has been incorporated with a deep and extensive expertise in fires behavior.

One of the most interesting adaptations is the Australian system, AIIMS (for Australasian Inter-Service Incident Management System), more used as Command Control and Coordination system (C3S), in powerful headquarters, as example in New South Wales state.

In the U.S. the FEMA (Federal Emergency Management Agency, www.fema.gov) is responsible, through its National Incident Management System, for organizing, teaching and training of this doctrine. In general, the position of this agency towards teaching and giving public information can be an inspiration at the European level.

The World Health Organization (WHO) has also set up an ICS-related system, WHE, (WHO Health Emergencies, 2016). WHE is a very close system, not allowing for much sectoral adaptation (quite copy and paste). Some industrial companies, especially in oil and gas, have also directly adapted ICS for their crisis management, for example Total Offshore and LyondellBasell (U.S. National Incident Management System presentation, 2017).

UN Organization

The United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA) framework United (Field Handbook, 2018) has, from its creation, been inspired as well by ICS. Although its field command post is called "On-Site Operational Control Centre" (OSOCC), it is not a real command and control system, but instead highly inspired by humanitarian activities. The main theme in UN field organizations is coordination (Figure 2.). So UNOCHA, is now in charge of managing a large certification system (International Search and Rescue Advisory Groups - INSARAG), and responsible for international aid, that is more focused on search and rescue and humanitarian aid.

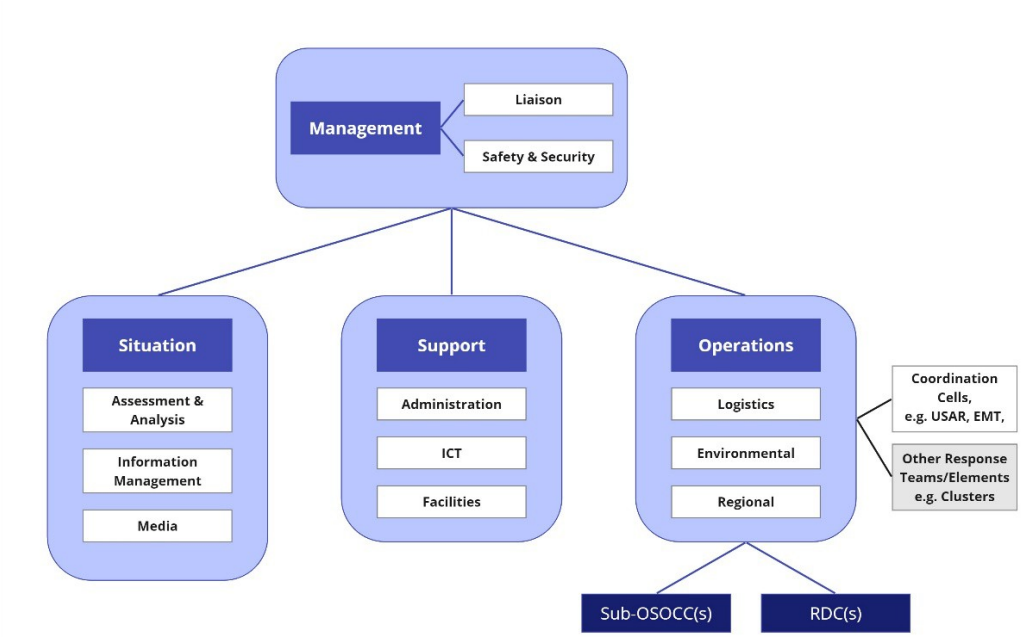


Figure 2. UN coordination

The OSOCC is a rapid response tool that provides a platform for the coordination of international response activities in the immediate aftermath of a sudden-onset emergency or a rapid change in a complex emergency. It is at the same time both a methodology and a physical location for on-site emergency response coordination. The OSOCC is designed to work in support of the Government of the affected country and is a tool promoted and supported by the UNOCHA to carry out its mandate of coordination and information management in emergency response, particularly at the field level.

Through its guidelines, the OSOCC enables the concept to be utilized by other organizations, agencies and teams, when responding to emergencies, including Governments, international and regional response organizations. It is a concept that can be internalized entirely by national emergency management structures who are in the lead of coordinating the assistance to populations affected by emergencies in their countries.

The OSOCC Guidelines are designed to provide direction on establishing, managing, working within and interacting with an OSOCC. The Guidelines are composed of three parts. Part I provides the foundation on which the OSOCC concept is built and the humanitarian context in which it operates. Part II presents guidance on how to implement the OSOCC system components, i.e., the Virtual OSOCC (VOSOCC), the Reception Departure Centre (RDC), the OSOCC facility, sub-OSOCCs and other coordination cells, and provides links to other more detailed guidance notes and resources. Part III is an annex of tools for practical use during OSOCC operations

The OSOCC is generally structured into four functions, each of which may be composed of multiple cells. The basic OSOCC structure is illustrated below; however, not all functions or cells may be needed in every situation.

French System, GOC.

In France, the GOC (standing for Gestion Opérationelle et Commandement) system has provided a unique and harmonized command system since the 1990's.
It manages four levels of command based upon progressive field transfers of command. Training program comprises five levels: level 1 is taught to all firefighters, it's not a commanding level. Level 2 (truck leader) is the first level of incident command used to deal with the majority (more than 50 %) of French fire service missions, tackled with a single resource. Levels 3 to 5 employ several ranks of officers, from team (group) leader to senior incident commander. Deployment of an Incident Command Post (ICP) starts at Level 4. Built as a decision tool for the incident commander, the ICP is staffed (one staff chief, four sections: operations, planning, resources & logistics, intelligence) with up to 12 persons; 20 in an interagency context. In this type of incident management team, the Incident Commander has a direct command on Sector Chiefs, whose number is usually between two and four (see Figure 3).

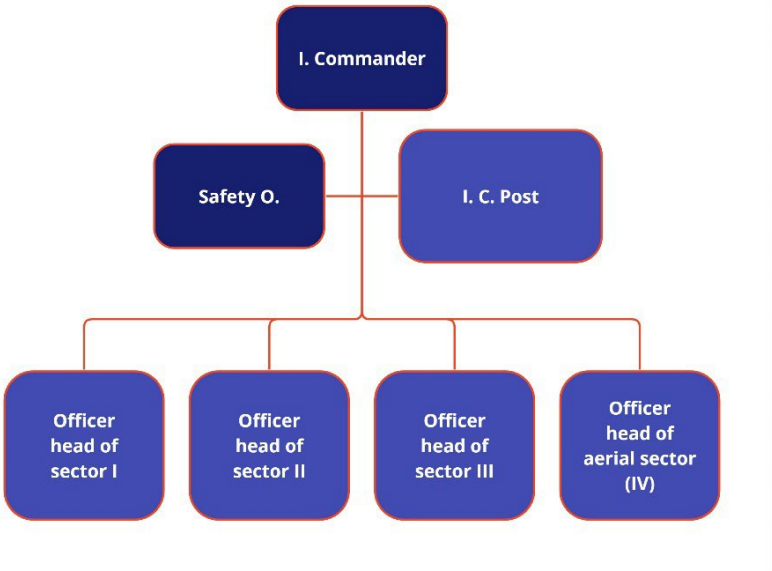


Figure 3. French GOC chart

This system has also been integrated by several countries of Africa, and Portugal and Spain (some provinces)

German System, DV 100.

Rather similar is the German system described in the set of regulations "Dienstvorschrift (DV) 100 - Führung und Leitung im Einsatz" (Service Regulation 100 - Command and Control in Operations) developed in the late 1970's after a major wildland fire and modified in the late 1990's. It combines the principle of sectorization with an incident command organized into functional sections. In most German states, an administrative crisis committee meets in parallel to the operational incident command in the event of major emergencies, both under the auspices of a political decision maker.

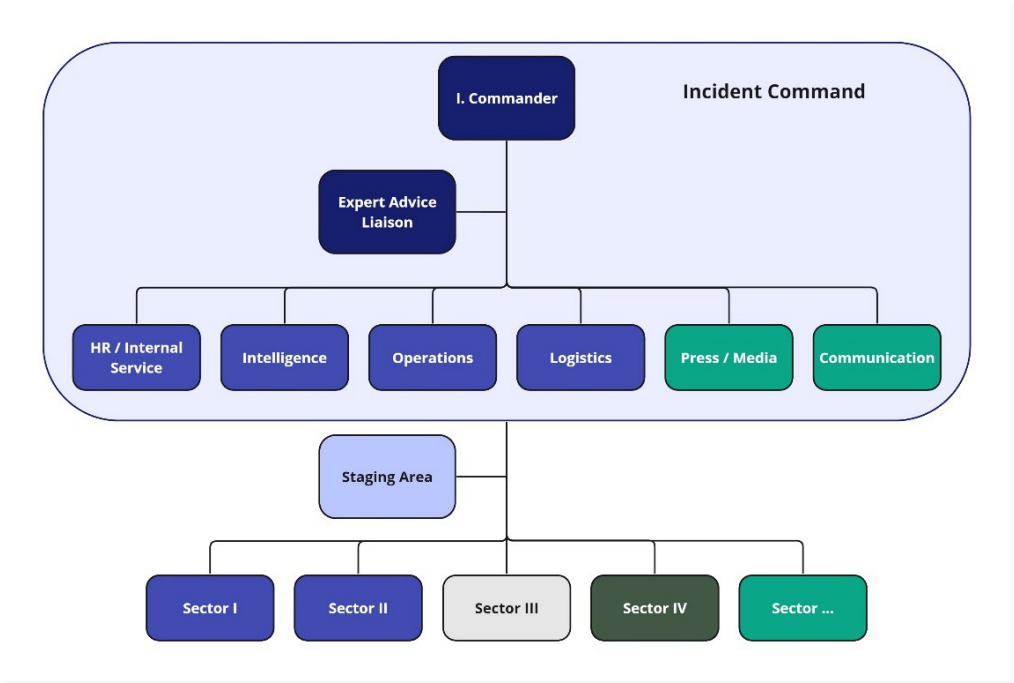


Figure 4. German GOC chart

NATO Framework.

Due to "*centric operations, demanding closer cooperation among more dispersed forces, the importance of interoperable C2 grows exponentially. Forces that expect to operate together must at least be able to communicate with each other via both voice and data formats, even though they are not yet equipped.*" (Barry, 2003). Indeed, for years, command and control has been considered by NATO strategists as the "*most crucial medium for interoperability*". Eventually, NATO shifted in 2017 toward a new command and control structure, but it's interesting to note that this structure is based on a constellation of separate international headquarters split in the different member states.

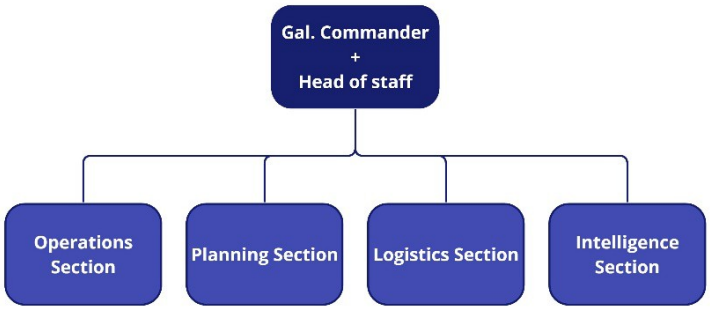


Figure 5. NATO basic chart

In the field, NATO is working with a very classical organization, close to a generic ICS. The four main functions (Figure 4.) are represented here in the scheme above: operations, planning, intelligence, and logistics. In many NATO operations, planning is done far from the field of operations, in a distant and secure headquarters.

The cited case in Sweden (Figure 5. below, 2018) provided a very interesting implementation of this framework, managed by a large incident management team (90 – 100 officers).

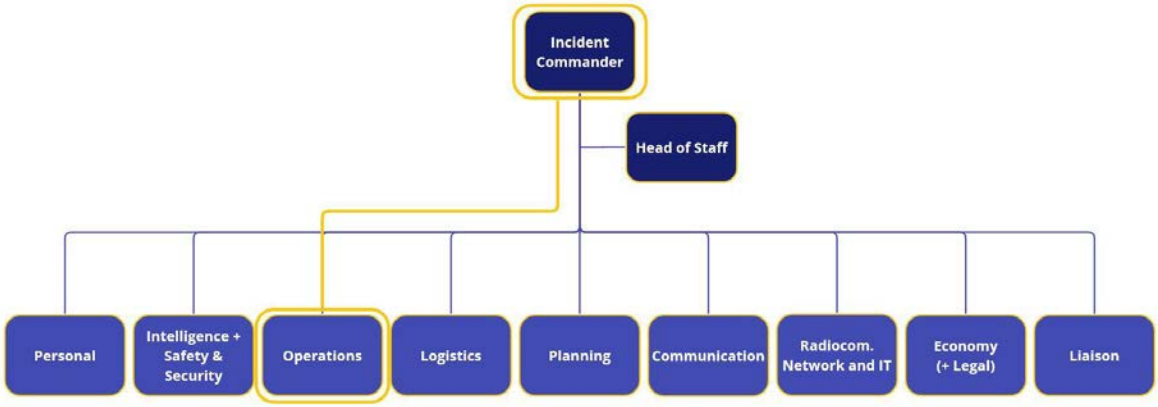


Figure 6. NATO / Swedish chart

After this quick overview of the existing landscape, two main concerns obviously appear: which relevant elements from these systems can be transferred to EUCPM with added value? And how to harmonize?

COMPARISON AND FINDINGS

To answer to the problematic mentioned at the end of the third chapter, we identified three critical issues:

- To manage very mature but not harmonized national civil protection systems, in all participant states;
- To be interoperable in and through the member states, some of them very small, respecting existing systems and implying a transition in the format of incident management when international aid is asked;
- Being compatible with UNOCHA teams and INSARAG standards, and others systems ICS, NATO, WHO and ISO 22320 compliant. But, fully independent from other international C2 systems, which can be sometimes considered as using “hegemonic practices”.

The idea developed here is not to copy and paste any system, but to develop, organize, implement and teach at EU level an ad hoc command system, ECS, tailor made for EU nations and disasters. It was foreseen since our 2018 works, but now somehow consolidated by our experimental training and new studies.

Comparison on the General Organization of the Command Framework

In all observed systems, the position of IC (incident commander) and its hierarchic relationship to LEMA plays a pivotal role. However, in some countries (Croatia for example), there is no clear differentiation between these two positions.

The connection with sectors or branches (generally 3) is sometimes done in a variable manner, but the recognition of a sectorization of the incident field is permanent and mandatory. Even if very dependent on national aerial laws and standards, the aerial sector appears to be fundamental for safety issues, in case of EUCPM activation. This sector is not developed in all countries' systems (only in GOC and ICS).

The existence of an Incident Command Post (or Field Command Post) is consolidated. Its internal organization is variable, but the common sections of quite all systems are:

- Operation
- Planning
- Resources

Alternatively, we find in the ICPs optional ones:

- Communication
- Intelligence
- Security
- Safety
- Coordination/liaison
- Finance-administration.

Safety and communication, frequently managed by single person positions besides the IC are also very dependent on command-and-control national culture. If these two functions seem to be a point of convergence, regardless of the nation, they are probably on the verge of turning into a function rather than a single officer duty. Other observed single positions (not ICP section) are: security (police), IT and radiocommunication, liaison officers.

Findings and Proposal

For large incidents, and activation of international European aid, the added value of this EU system could be in providing and establishing an interoperable framework of art of command, adaptable at state and at European level, considering national command systems and history.

This is the reason why we consider as “strongly suggested” the sections: **operations, logistics and planning**.

The single positions of **safety officer and communication** are at least recommended.

The **sectorization**, either geographical or missional, and aerial sector if the operation encompasses aircrafts, is also our recommendation.

Eventually, we identified a common “backbone” of command-and-control functions (Figure 6.)

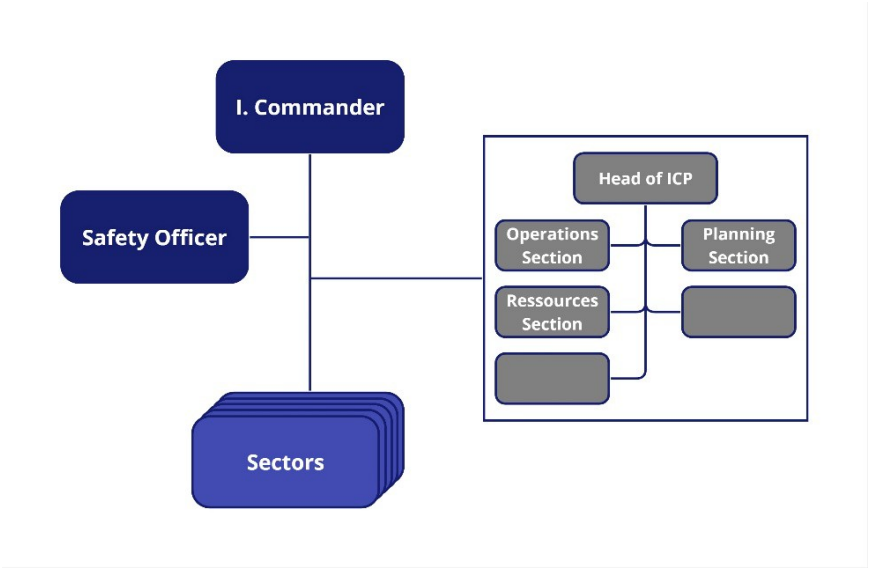


Figure 7. Commonalities

The proposed European Command System is, in evidence, a long-term solution. It must be expanded upon from this skeleton onwards. Rather than standardization, DG ECHO might establish guidance and recommendation to be acknowledged by practitioners.

Additionally, it is highlighted that in this way of construction the proposal is ISO 22-320 compatible, fulfilling the requirements listed at the beginning of the section.

What is nowadays lacking is a reliable inventory of all European participant states' command and control systems. That could support the authors to validate their study.

SHARING KNOWLEDGE

Since 2001 DG ECHO has organized EUCPM training for Member States. Renewed at the end of 2023, and reinforced by “knowledge network” website, this framework is now enriched by many courses. It also distinguishes clearly between the courses for responders and those dedicated to civil servants or staff persons.

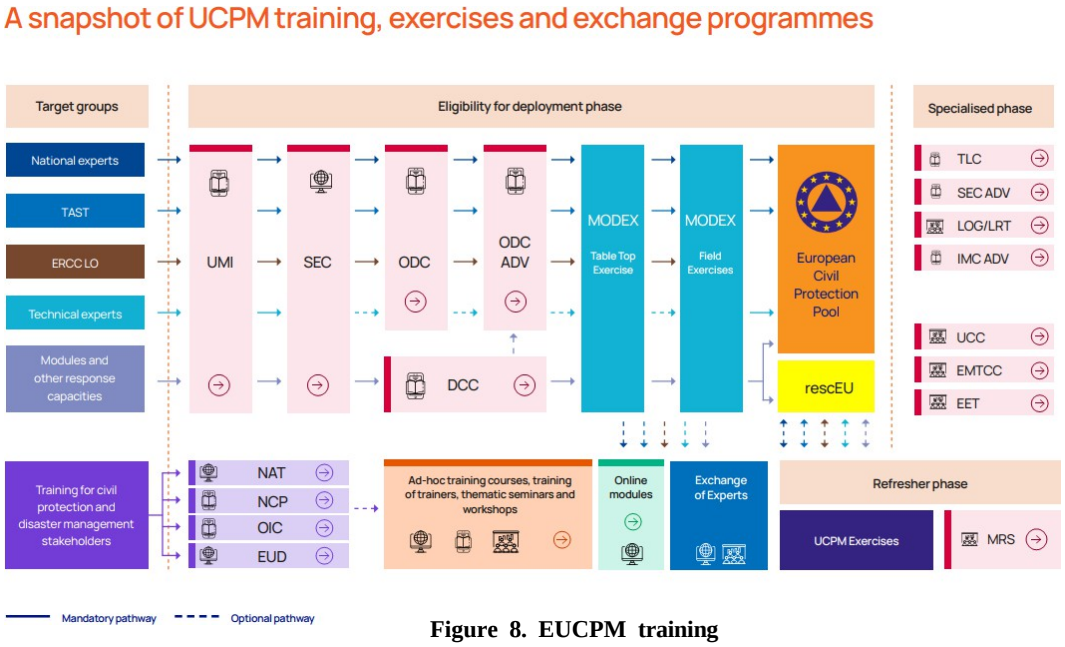


Figure 8. EUCPM training

The authors' suggestion is that DG ECHO would need to set, manage, and teach a "new command and control framework", or at least share the knowledge of the international C2 systems used within Europe, to increase our civil protection interoperability.

Finally, the authors propose to simply use the already existing training program (EU Brochure, 2023 & fig.7), and link these proposed trainings (see below) with On site Deployment Courses (ODC in the scheme).

Experimental Trainings

Several experimental European courses were organized in Aix-en-Provence, at national wildfire school (ECASC) France, in the context of European project called *"Network of European hubs for civil protection and crisis management"* (Ecorys, 2020), and in the Green deal project FIRE-RES (April 2024).
Gathering 40 trainees from 20 different countries, these trainings were organized at the French fire service facility's virtual reality simulator, involving a complete incident management team (15 to 20 players). This serious game provided a unique opportunity to organize some trials and experiments, as described below.

Methodology:

During experimental courses we highlighted that at least in Europe, command and control systems are comparable and some common points may be underlined.

The aim of this experimental training was to establish and drive an incident management team on a big incident and to deal with the arrival of – at least – two EU modules.

Exercise Incident Command and Control: The management of the training proposed a fixed command framework (IC + 3-4 sectors + ICP, cf. Figure 6.). But otherwise, the organization and staffing of the Incident Command Post was fully adaptable to each Incident Commander (i.e. to each nation receiving UCPM aid).

We only had a small "sample" of countries to lead this Incident Command System simulation: Italy, France, Sweden, Croatia, Portugal and Greece. However, all attendees were involved in ICP functions or single officer positions (safety, public info. officer, liaison officer …), giving a unique opportunity to test full international (Austria, Belgium, Poland, Croatia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Netherlands, Norway, Portugal, Slovakia, Spain, Sweden) interoperability. This experiment demonstrated that EU participant states are able to cooperate efficiently in a command-and-control system. It also

confirmed the incident management commonalities existing between all participant states. This common framework could obviously constitute the backbone of a future European command and control system.

Modules and Experts Involvement:

This structured, common command and control system is actually very helpful to optimize involvement of EU modules. To manage module national specificities and logistics, a very obvious methodology (more than 9 times tested in the experimental courses) emerged at the end of these trainings: the Chief of Module needs to be directly involved within a sector, while a Liaison Officer needs to stay assigned to the ICP, in order to be the permanent representative of the module (logistics, capabilities, safety, etc …). This best practice is now beginning to be shared at EU level, even if it is not a mandatory procedure. Additionally, testing international cooperation within a unique incident management team in real/simulated conditions proved its feasibility and constitutes a very promising input for the future embedding of EU experts into the host nation incident management team.

This small-scale experience, more empirical than scientific, constitutes a demonstration of our theory on this parcel of interoperability and gives arguments and a possible form for future EU trainings.

CONCLUSION

We have imagined a doctrinal framework, totally dedicated to civil protection, based on guidelines developed in this paper. Effective, interoperable, respectful of sovereignties, this system could match with the needs and wishes of the EU for its citizens, and promote solidarity between the EU and non-European regions.

Of course, this proposal is anchored in experience rather than on scientific research. Further studies could be performed to confirm and consolidate this proposal. What is nowadays lacking is a reliable inventory of all European participant states' command and control systems. That could help to validate our assumption on the “common backbone”, and to clearly differentiate civil protection operations from humanitarian ones.

This work faces a paradoxical time context: after being hampered by Covid, its immediacy has been triggered by the climate change effect since 2021. Regarding usual experience with instructing, training and exercising efforts, it could take years to begin to fully use the new approach in operational situations. The first evaluation feedback is foreseen to be available after five years.

The European Commission has to determine how to organize the development of an improved EUCPM. How to concern and involve national civil protection, fire and EM services structures? What will be the role of national Fire and civil protection academies? How to let this European Command System knowledge percolate towards national levels? All these key points need to be solved if we really want to improve the interoperability of the participant states and the efficiency of EUCPM, to give European citizens the safety they deserve.

GLOSSARY

C2 Command and control

C2S Command and control system

ECASC ECole d'Application de la Sécurité Civile (Valabre national wilfire academy, France)

ECS European Command System

ERCC Emergency Response and coordination Center

EUCPM European Union Civil Protection Mechanism

ENSOSP Ecole Nationale des officiers de sapeurs-pompiers (French national officers' Fire academy)

EWE Extreme Wildfires Events

FCP Field Command post

GOC Gestion opérationnelle et commandement (Command and Operational management)

ICP Incident Command Post

ICS Incident command system, generic term which is quite a trademark US system

INSARAG International Search And Rescue Advisory Group

LEMA: Local Emergency Management Authority, Decisional policy authority in incident management

NGO, NGOs non-governmental organization(s)

OSOCC: On Site Operations Coordination Centre

STPs Standard Technical Procedures

SCCPs Standard Command and Control Procedures

SOPs Standard Operating Procedures

UNOCHA United Nations Office for the Coordination of Humanitarian Affairs

UNDAC United Nation Disaster Assessment and Coordination

REFERENCES

Barry, L. C. (2003) Transforming NATO Command and Control for Future Missions, *Defense Horizons issue #28,* Center for Technology and National Security Policy National Defense University, Washington DC.

Brisset, J-V. (2006) Europe et Nrbc : quel est l'état de la coordination européenne en cas d'attentat terroriste Nrbc ? Quelles perspectives ? Iris, Paris FR.

Dufès, E. and Ratinaud, C. (2014) Situations de crise : une réponse modélisée en 3D, *Perspective n°12, p. 57-78,* Ensosp, Aix en Provence FR.

Ecorys (2020), *Network of European hubs for civil protection and crisis management,* Final project report for DG ECHO, Rotterdam NL.

Essen, M., McCaffrey, S., Abrams, J. & Paveglio, T. (2022). Improving wildfire management outcomes: shifting the paradigm of wildfire from simple to complex risk, *Journal of Environmental Planning and Management, 66:5, 909-927,* Tailor & Francis online publishing, London, UK, DOI:10.1080/09640568.2021.2007861.

European Commission (2016), *The Union Civil Protection Mechanism Training Programme*, booklet, DG ECHO, Luxembourg LU.

Greco, P. (2017) Lyondell and U.S. National Incident Management System (NIMS) Introduction and Overview, *Proceedings of EAI – Proceedings of Emergency Response & Fire Protection Conference,* Houston TX.

Maestracci, B. (2013) La gestion de crise, perspective européenne, *Perspectives n°9*, p. 159-161, Ensosp, Aix en Provence FR.

Monet, J.P, Joseph, X., Robertson, T., Saint-Jonsson, A., Mannaioni, A., Schaller, P. (2023). D4.8 Matricial Tool and Ad Hoc Guidance. *Deliverable 4.8 FIRE-RES Eu funded project.* 15 https://doi.org/10.5281/zenodo.8377362 https://fire-res.eu/wp-content/uploads/2023/11/D4.8_FIRE-RES_Matricial_Tool.pdf

San-Miguel-Ayanz J., Durrant, T., Boca, R. et al., (2023) *Forest fires in Europe, Middle East and North Africa 2022,* European Commission, Joint Research Centre report, Publications Office of the European Union, Ispra IT-Brussels B, https://data.europa.eu/doi/10.2760/348120

Tedim, F., Leone, V., Amraoui, M., Bouillon, C., Coughlan, M.R, Delogu, G.M., Fernandes, P.M., Ferreira, C., McCaffrey, S., McGee, T.K., Parente, J., Paton, D., Pereira, M.G, Ribeiro, L.M., Viegas, D.X. and Xanthopoulos, G. (2018). Defining Extreme Wildfire Events: Difficulties, Challenges, and Impacts. *Fire 2018, 1, 9,* MDPI Publishers. Basel CH https://doi.org/10.3390/fire1010009

Topper, B. and Lagadec, P. (2013) Fractal Crises - A new path for crisis theory and management, *Exploring the Theoretical Foundations of Crisis Management of Journal of Contingencies and Crisis Management, special Issue,* Wiley-Blackwell; Oxford UK.

UN Office for the coordination of Humanitarian Affairs (2018), *United Nations disaster assessment and coordination Field Handbook*, 7th edition, Geneva CH.

World Health Organization (2016), *Update Who Health Emergencies Programme: progress and priorities; Financing dialogue*, Geneva CH.

Wybo, J.L. Latiers, M. (2006) Exploring complex emergency situations' dynamic: theoretical, epistemological and methodological proposals, *International Journal of Emergency Management,* Inderscience Publishers, Geneva CH.

Unified command? Preliminary findings from a situation awareness experiment

Jeroen Wolbers
Leiden University
j.j.wolbers@fgga.leidenuniv.nl

ABSTRACT

This study investigates high-pressure decision-making in fire response operations, with a particular focus on the cognitive abilities of incident commanders. It builds upon the naturalistic decision-making approach, emphasizing the pivotal role of situational awareness in decision-making. While existing studies predominantly concentrate on individual decision-makers, this study shifts attention to the collaborative dynamics within incident command teams. Through computer-based experiments, it explores the impact of interdependence between fireground commanders on situational awareness, challenging the presumed benefits of teamwork. Counterintuitive results question the conventional wisdom regarding the advantages of having a second pair of eyes. The experiment's findings indicate that commanders operating alone achieve the highest situational awareness scores, suggesting that organizing periodic moments of consultation works better than operating in pairs during response operations. This study provides insights into the intricate interplay of interdependence and situational awareness during high-pressured decision-making.

Keywords

Situational awareness, decision-making, command & control, experiments, fire department.

INTRODUCTION

Try to imagine a fire engine rushing to the incident scene, blue light flashes, sirens, and rapid commands uttered by the crew commander to his team. While squeezing through heavy traffic, crews put on their gear, getting ready to battle a raging fire. Probably one of the key concerns on the mind of the engine commander and his crew is getting a bearing of what they will run into. A sense of awareness and vigilance. Not only will this aid in the rapid rescue of victims, it is also imperative to safeguard the lives of the crew. When the fire crew arrives on-scene and rushes onto the fire ground, they will need to carry out a rapid risk assessment and take critical decisions that will likely determine the difference between life and death.

Such tense situations of high-pressured decision-making have attracted the attention of academics. The classic work of Klein et al. (1986; 2010) on rapid decision-making on the fireground marked the beginning of a fruitful line of research. The study departed from the classic idea of rational, stepwise decision-making and instead stressed the key role of experience by which decision-makers interpret regularities in the environment. Traditional decision models did not take into account the critical role of experience to assess operational settings, as they often relied on students participating in decision-making experiments. Instead, Klein et al. (1986) forwarded a naturalistic approach, in which decision-making is closely connected to the environment, in which previous experiences and expertise plays a major role in guiding the decision-making. It was found that experts are able to read regularities in the environment on which they develop heuristics or rules of thumb, which function as cognitive shortcuts that enable rapid decision-making based on recognition (Klein et al., 2010).

As the naturalistic decision-making school gained footing, many studies on incident command teams followed suit, focusing on the cognitive abilities of the commander (Boehm, 2018), such as situational awareness (Endsley, 1995), sensemaking (Barton et al., 2015; Barton & Sutcliffe, 2009), intuition and mental simulation (Klein et al., 2010). Other studies documented the sequential order of decision processes (Cohen-Hatton et al., 2015; Groenendaal & Helsloot, 2016). This resulted in a renewed understanding of the decision-making process, in which a different sequence of decision phases is identified, which departs from the logical stepwise order seen in normative models (Cohen-Hatton et al., 2015). In these models phases like situation assessment, plan formation, and plan execution are identified as important phases in the decision process (Cohen-Hatton et al., 2015). Initial

results support the idea that in high-(time) pressured situations decision-making is primary guided by expertise as decision-makers skip plan formulation. It is also found that goal-directed training can aid decision-makers to switch to more deliberate refection in action (Cohen-Hatton et al., 2015). Taking the time to question perspectives and renewing one's situational awareness remains a key challenge in time pressured environments.

If one thing stands out from this line of research is that most studies focus on the cognitive abilities of the individual decision-maker. Given the attractiveness of the role of the commander in this setting, this is understandable, but also is a limitation that warrants our attention. There are important empirical and theoretical limitations to this focus. A key concern is that the individual decision-maker is often biased and is prone to make intuitive errors (Kahnemann & Klein, 2009). Moreover, a range of studies point to the importance of *shared* and *distributed* sensemaking (Barton, et al., 2015; Wolbers, 2022), *collective mind* (Weick & Roberts, 1993), and *dynamic delegation* (Klein et al., 2006). Indeed, in many operations a commander is not alone, but part of a hierarchy with multiple lines of command. As incidents turn more complex, multiple commanders work in parallel and interact with each other. Therefore, it is time to take distributed command into account more elaborately.

This study explores the impact of collaboration between commanders on situational awareness. Does it help to organize backup for a commander that is operating in a dynamic and chaotic scene? The question is posed: *what is the influence of different types of interdependence between fireground commanders on situational awareness in a fire response operation?* In order to answer this question, the concepts of situational awareness and interdependence are discussed and later operationalized in the methodology. The design and execution of the computer-based experiments offers an innovative way to study the role of teamwork in command and control studies. This study offers counter-intuitive results that question the benefits of teamwork for reaching a high level of situational awareness.

FROM SITUATIONAL AWARENESS TO SYNCHRONIZED DECISION-MAKING

Command & control is a key process in incident response operations that determines what actions response organizations need to take. There are multiple studies that map what incident commanders do when they take command (Njå & Rake, 2009). These studies show that incident commanders usually focus on problem solving and are operating within narrow time horizons. Njå & Rake (2009) note that command strategies were often reactive and interestingly, that very few commanders turn to giving commands. Most of the incident commanders were found to give very few commands (Groenendaal & Helsloot, 2016). Instead, commanders are busy with developing situational awareness in order to monitor the work of others and intervene only when absolutely necessary.

Studies into incident command tend to use a naturalistic decision-making approach (NDM). NDM centers its focus on the proficient decision maker, who operates within bounded rationality. The NDM framework comprises two primary components: situation assessment and mental simulation, both of which are context-specific and rooted in experience. The RPD model, as proposed by Klein et al. (1986), draws upon research spanning diverse domains, including (wildland) fireground command, military battle planning, critical care nursing, ship and flight deck operations, and chess tournaments (Njå & Rake, 2009). In essence, the RPD model depicts decision-making as a sequential process, where proficient decision makers assess the situation and choose an action based on their expertise. The model involves continuous monitoring of deployed actions, and incorporates feedback loops to account for the effects of actions taken. Additionally, mental simulations in the RPD model facilitate feed-forward loops, enabling a proactive approach to decision making.

In contrast to the naturalistic decision-making (NDM) approach, the heuristics and biases (HB) approach points out that humans are often inconsistency and reach different conclusions based on similar information (Tversky & Kahneman, 1974). Tversky & Kahnemann (1974) argued that human judgment is often flawed due to unconscious bias. The naturalistic and HB are often pivoted against each other in terms of influence of expertise. When integrating the insights from both approaches, Kahnemann & Klein (2009) conclude that the quality of intuitive judgment relies on the predictability of the environment and the individual's opportunity to learn the regularities of that environment. Both agreed that situational awareness formed the ground on which decisions were made (Kahnemann & Klein, 2009).

Assessing the environment thus plays an important role in decision-making processes, which points to a key role for situational awareness. This is corroborated by studies that investigate the sequential order of decision-making process, who point out that the step of situation assessment is leading in decision processes (Cohen-Hatton et al., 2015). According to Endsley (1995: 36), *'situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future'*. This definition identifies three levels of situational awareness: (1) perception of elements in the environment, (2) comprehension of the meaning of those elements, (3) and the projection of a future state.

Whereas most often situational awareness is frequently measured on the individual level, a range of studies has explored team-level dynamics (Artman & Garbis, 1998). These studies point out that teams play a key role the quality and direction of decision-making (Salas et al., 2017). For instance, Stanton et al. (2017) discuss the role of distributed awareness in dynamic systems and explain that each actor holds their own situational awareness, which may be very different from others. While large disparities may emerge between actors, they note that it is not always necessary to share a similar level of situation awareness, as commanders may have different roles or focus. In that sense, team situational awareness may be understood as a dynamic, fluctuating process, binding agents together on a moment-by-moment basis (Treurniet & Wolbers, 2021). Similarly, Salas et al. (1995) argue that team situational awareness is comprised of a cyclic process that oscillates between individual and team level awareness. On a slightly different note, Shu & Furuta (2005) point out that the required awareness depends on the intensity of a collaboration, and introduce the concept of mutual awareness. Mutual awareness describes the situation where two actors interact intensively for the completion of a task. Therefore, in order to understand the required level of situation awareness, it is necessary to look into the intensity of different forms of collaboration.

In order to find out whether different forms of collaboration impact the level of situational awareness, the concept of '*interdependence'* is helpful. This concept from organization theory lays the foundation for what type of coordination and sensemaking can be achieved in organizing processes (Weick, 1993; Wolbers & Boersma, 2013; Wolbers et al., 2018). Thompson (1967) identifies three forms of interdependence: *pooled*, *sequential* and *reciprocal*. Pooled interdependence refers to arrangements were a task is split among actors that work independently with no flow of work between them, requiring little contact. Sequential interdependence occurs when actors rely on one-directional pathways that contain predictable workflows, which are often codified into plans. The idea of a conveyer belt in a factory comes to mind, to describe a situation where one step is required in order to a colleague to take a next step. Finally, reciprocal interdependence describes the situation where actors work back and forth and mutually adjust their actions.

These three types of interdependence are crucial in the setting of crisis management, as they describe how the composition of, and interaction between command teams may influence distributed sensemaking (Schakel & Wolbers, 2021; Wolbers, 2022). How mutual interdependencies influence cognition is described by Weick et al. (2005), who notes how each type of interdependence is tied to a specific type of cognition. Pooled interdependence is associated with automatic cognition through stereotypes, as actors are not in the position to communicate directly, but have a share sense of how a task should be completed. Discussions on collective mind also describe such a type of meta-cognition (Weick & Roberts, 1993). Sequential interdependence is often tied to heuristic cognition associated with memorized if–then rules, which builds upon the sequential order between work processes. We can recognize this type of recognition in the naturalistic decision-making school, discussing heuristics as a way to codify the expertise of decision-makers in dynamic situations. Reciprocal interdependence involves a more controlled type of cognition based on doubt, inquiry, and argumentation, which emerges from the continuous interaction between decision-makers. Indeed, reciprocal interdependence is time intensive, but also important as actors are able to question and doubt dominant frame that develop in a response operation (Barton et al., 2015). By engaging in this type of controlled cognition actors are able challenge existing frames and able to prevent a collapse of sensemaking (Weick, 1993). These studies show the key role that the different types of interdependence may play in influencing the cognitive abilities of incident commanders.

METHODS

In this study, we measured the impact of different types of interdependence (pooled, sequential and reciprocal) on the level of situational awareness of fire ground commanders. We have created a 1,5h scenario in a virtual reality environment of a complex fire in a storage facility that commanders were asked to watch. The commanders cannot walk around in the computer simulated environment themselves, but have to watch 5 extracted video fragments in which the incident response simulation unfolded, in order to show each commander exactly the same information. The commanders are asked to answer 15 probe questions after each video in order to measure their level of situational awareness. In order to determine the effect of the type of interdependence, we have randomly assigned the commanders into three corresponding groups. We designed these manipulations according to the concepts of pooled, sequential, and reciprocal interdependence (Thomson, 1967). The control group engaged in pooled interdependence, which meant that subjects viewed the videos completely independent (alone). In the manipulation of sequential interdependence, subjects were asked to work together, but were only able to update each other after watching a video, just before answering the questions. In the manipulation representing reciprocal interdependence, the subjects were allowed to engage in full, continuous consultation while watching the videos. In this way, we were able to simulate the three different types of interdependence in the experiment.

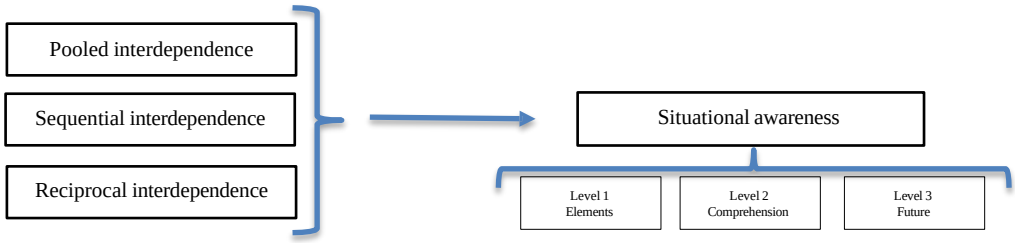


Figure 1. Conceptual design of the experiment

We measured situational awareness according to the QATA methodology, as described in Edgar et al. (2018). In this method, probes are presented to the subjects in the form of visual cues, situation reports, or radio messages. Subjects are presented with relevant and irrelevant probes during the videos, and depending on the manipulation group they are in, are able to discuss their assessment with another commander. The QATA methodology uses a confirm or deny format, as subjects can select that questions are true or false. To measure situational awareness, we formulated 75 probe questions. Each video contained 15 questions, 5 corresponding with each level of situational awareness: level 1) perception of elements, level 2) comprehension of the coherence between the elements, and level 3) future projection. We used the same set of probes and questions throughout the entire incident across all subjects. Subject were asked to state whether a probe was true or false, and how confident they were about each answer on a 5-point Likert scale. Examples of the probe questions are: *'there are people in front of the print shop that make access to the launderette difficult'* [level 1]; '*since the color of the smoke has changed, the situation in the Laundromat is deteriorating'* [level 2], '*heat affected IBCs behind the building pose a risk for a rapidly developing fire'* [level 3].

Participants

112 fire commanders volunteered for this study and provided informed consent for their participation. They were drawn from a representative sample covering all 25 safety regions in the Netherlands. The sample included 103 males and 9 females. The average age was 45 years. The participants were randomly assigned in three groups on the basis of whether they were to engage in pooled, sequential or reciprocal interdependence during the experiment. The participants were not informed there were three different groups. Each participant was briefed on the procedure and gave informed consent for their participation.

We determined the minimal number of participants in advance, by conducting a G-power effect size calculation. To find a large effect size, we required 21 participants in each group, with 3 groups manipulations, this means 63 participants. To find a medium effect size, we required 52 participants in each group, with 3 group manipulations, so 156 participants. We concluded that we needed a sample of at least 52 participants up to 156 participants to measure a medium effect size. In total, 112 commanders participated in the experiment, with 44 in pooled, 34 in the sequential group, and 34 in the reciprocal group. Previous operational experience was assessed in a questionnaire leading up to the experiment that identified how long participants had served in the role of on-scene commander. The mean overall command experience was 9.33 years.

Equipment

The experiment was designed by making use of the software XVR Simulation. The XVR Platform provides all fire departments in the Netherlands with computed-based scenario training. Trainees are immersed in a controllable situation, which is repeatable and measurable. Similar research with support of simulations exists, such as C3 fire (Johansson et al., 2003) and FireChief (Omodei & Wearing, 1995). XVR is step forward in this sense, as it provides the subject with an immersive realistic 3D view of the incident scenario. This paper builds on successful experiences with XVR software in the Firemind and FireFront research projects (Thoelen et al., 2020; Polikarpus et al., 2022). The scenario was designed from scratch, for which we made use of fire department experts, and scenario trainers in a safety region in the Netherlands. We consulted with these experts to make sure to develop a scenario that is realistic for fire-fighters, which contain cues that they would also encounter in a real-live situation, such as the of the spread and intensity of the fire, color of the smoke, fire hoses, fire engines, and other emergency services. We created a 5 separate videos from the 1st person perspective of our avatar walking through the incident, as the scenario unfolded.

The subjects watched the videos on the Labvanced platform, which we used to build online experiments. The platform was used to build, record and share online experiments directly in the browser. It also enabled multi-user studies, in which two subjects perform a task together and the system waits for them to complete the task together before moving forward. In this way, we made sure participants watched the same videos and answered the questions at the same time. We asked the participants to conduct the experiment on their home desktop, in order to provide a calm environment. The subjects in the sequential and reciprocal interdependence groups stayed in

contact with each other through a video call on Microsoft Teams in a separate desktop window, which enabled them to interact with each other throughout the scenario. This channel was recorded and continuously monitored by two research assistants to check for disturbances or technological problems. Before the experiment was launched, two research assistants checked in with respondents to check the setup, ensure a quiet environment, and explain the logic of the experiment based on a predefined procedure with fixed talking points. As a result of this monitoring, the data of one participant was deleted due to technical problems with the microphone.

It is important to recognize a few limitations of this setup as the commanders cannot 'seek' information as they normally would do during an incident as they watch pre-recorded videos, are not inside an operational setting itself, and do not continuously see their real-life counterparts. This was accounted for in the setup by showing prerecorded conversations with virtual role players, and information that reached commanders through visual cues, pager notifications, and radio messages. The advantage of this setup is that we were able to create exactly the same operational environment for all participants, the downside is that participants could not adjust it.

Nature of the Scenario

The scenario depicts a complex fire in an industrial warehouse building with several compartments, housing a laundry, garage, car repair shop, and a printshop. The fire starts in the garage, but soon spreads to the laundry and car repair shop. As the fire and smoke grow in intensity, a hidden XTC-lab in the back of the laundry catches fire and explodes. A fire fighter is reported missing and part of the building collapses. In this way, the scenario starts relatively simple, but as the videos progress increases in terms of complexity. Cues are offered to the subjects that hint to an XTC lab inside one of the compartments, before and after the collapse, such as reports of interference with electrical wiring, blue barrels, and chemical equipment. There are multiple fire department processes that needs to be arranged simultaneously (asking for additional engines to provide water, scaling up the operation, and deciding on the entry tactics). In that way, we simulate a growing cognitive load for the commanders.

The first video covers the initial reconnaissance walk as the commanders arrive for the first time. It shows an industrial complex on fire, with smoke pillowing from multiple windows. The second video shows the initial attack by fire crews, in which there is an escalation of the fire/smoke propulsion, which leads up to the question whether or not to ventilate. The third video shows the situation directly after the explosion, in which the north side of the building has collapsed and possible persons missing. Here we also introduce the XTC barrels for the first time. The fourth video shows the XTC lab in detail and the fire escalating in other parts of the premises, so that the subjects have to choose between focusing on the XTC lab, or focusing on the fire that is developing in other parts of the building. At the end of the fifth video, the fire fighter that is reported missing is found and the fire is contained by the crews on site.



Screenshot 1. First view on the fire



Screenshot 2. Crews going into the laundry for reconnaissance



Screenshot 3. First cue of the toxic waste from the XTC lab



Screenshot 4. Situation after the explosion and collapse

Figure 2. Screenshots 1 - 4 from the videos shown to participants during the simulation

RESULTS

The first results of the experiment indicate significant differences between the pooled, sequential and reciprocal interdependence, both at the combined situational awareness (across levels), and in particular at the second SA level (comprehension). A score ranges from 0 - 100 and means that the subject has a good overview of that percentage of the cues. That means that a score of 100% is not realistic, as subjects are not expected to remember all cues, but discriminate important and less important cues. The score does give an impression of the scope of situational awareness, not so much whether crucial signals are missed or not. All groups appear to be relatively capable of building up an adequate level of situational awareness.

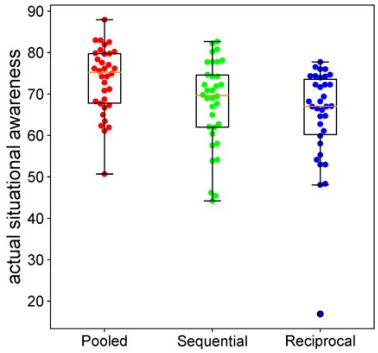


Figure 3. Actual Situational Awareness

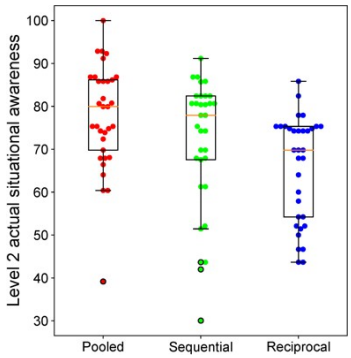


Figure 4. Level 2 ASA

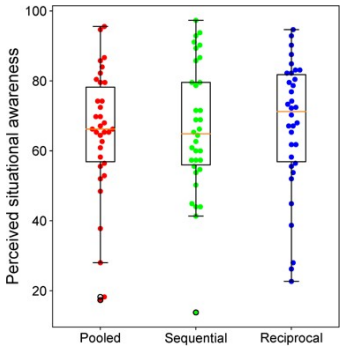


Figure 5. Perceived Situational Awareness

The analysis offers two important statistically significant results on the measure of overall actual situational awareness (ASA) and level 2 situational awareness (L2ASA). A Kruksal Wallis non-parametric test was conducted to test whether the medians of the different groups are equal. The Kruksal Wallis was found to be most fitting to the dataset, as it does not assume a normal distribution, in contrast to a regular one-way ANOVA. In a Kruksal Wallis test the is data reorganized based on a ranking.

For overall situational awareness (ASA), the Kruksal Wallis test showed that individual exploration of the scene significantly affects the level of situational awareness, compared to actors that engage in continuous collaboration: $H(2) = 20.92$, $p = .004$. Commanders exploring the scene alone based on pooled interdependence (*Mdn* = 73.26), or engaging in a form of sequential interdependence (*Mdn* = 69.64), scored significantly higher than commanders engaging in reciprocal interdependence (*Mdn* = 67.03).

For level 2 situational awareness (L2ASA) that focuses on the comprehension of the meaning of elements in a scenario (i.e. overview of the situation), results show that engaging in individual exploration significantly affects the level of situational awareness: $H(2) = 22.29$, $p = .002$. Here, commanders engaging in a form of sequential interdependence (*Mdn* = 76.63), score higher than in exploring the scene individually in pooled interdependence (*Mdn* = 75.33), and scored a significantly higher than commanders engaging in reciprocal interdependence (*Mdn* = 69.81).

These results indicate that the subjects in the pooled interdependence group had a higher situational awareness score than the subjects engaging in periodic or continuous consultation. In comparison, the subjects engaging in continuous consultation (reciprocal interdependence) have the lowest scores on average. In Figure 5, the perceived situational awareness (PSA) is indicated, of which the scores are not significant and are dispersed across the plot. This could suggest that subjects may find it difficult to make a valid assessment of their own situational awareness.

CONCLUSION

The results of this experiment indicate that when commanders engage in continuous conversation during a response operation, it limits their situational awareness. Commanders that take in cues on their own, score significantly higher than commanders working in teams. This is an important result that shows that having two pair of eyes does not always result in a better situational awareness.

It is important to note that this study concentrates on situational awareness, whereas this is not the only aspect at stake during a response operation. The required decisions, interaction with other crews and emergency services, may very well warrant the help of other commanders. It is unlikely that a commander will be in the position to develop situational awareness all by him/herself during the entire incident. These results of this study also provide valuable lessons for shared incident command. If a second commander is called towards the scene, he or she has the most positive contribution if he/she takes in the situation before coming into contact with the other commander. During a response operation, it works better to organize periodic moments of consultation, compared to operating in pairs. That makes it important to make process agreements in command teams on how commanders will collaborate, when to separate, and when meet up again. Although commanders are of course not able to spot all cues, engaging in continuous joint exploration is also not the solution. A good assignment for a second commander could be to focus on noticing abnormal signals that he or she feels have not been picked up by the other commander(s). By distributing tasks in this way, the mental load during high-intensity decision-making can be reduced. This makes the assignment of the second commander twofold: 1) mapping the situation independently of other commanders; 2) find out what the perception of the other commander is, in order to complement and challenge existing beliefs. As such, these findings have important implications for the development of the command tactics and situational awareness in situations of high-pressured decision-making.

REFERENCES

Artman, H., & Garbis, C. (1998). Situation Awareness as Distributed Cognition. *Proceedings from Proceedings of 9th Conference of Cognitive Ergonomics*, Republic of Ireland: Limerick.

Barton, M. A., & Sutcliffe, K. M. (2009). Overcoming dysfunctional momentum: Organizational safety as a social achievement. *Human Relations*, *62*(9), 1327–1356.

Barton, M. A., Sutcliffe, K. M., Vogus, T. J., & DeWitt, T. (2015). Performing under uncertainty: Contextualized engagement in wildland firefighting. *Journal of Contingencies and Crisis Management*, *23*(2), 74–83.

Boehm, M. (2018). Recalling the performativity of the body in frontline command. *Journal of Contingencies and Crisis Management*, *26*(4), 461-468.

Cohen-Hatton, S. R., & Honey, R. C. (2015). Goal-oriented training affects decision-making processes in virtual and simulated fire and rescue environments. *Journal of Experimental Psychology: Applied*, *21*(4), 395.

Cohen-Hatton, S. R., Butler, P. C., & Honey, R. C. (2015). An investigation of operational decision making in situ: Incident command in the UK fire and rescue service. *Human Factors*, *57*(5), 793-804.

Edgar, G. K., Catherwood, D., Baker, S., Sallis, G., Bertels, M., Edgar, H. E., ... & Whelan, A. (2018). Quantitative Analysis of Situation Awareness (QASA): modelling and measuring situation awareness using signal detection theory. *Ergonomics*, *61*(6), 762-777.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors*, *37*(1), 32-64.

Groenendaal, J., & Helsloot, I. (2016). A preliminary examination of command and control by incident commanders of Dutch fire services during real incidents. *Journal of Contingencies and Crisis Management*, *24*(1), 2–13.

Johansson, B., Persson, M., Granlund, R., & Mattsson, P. (2003). C3Fire in command and control research. *Cognition, Technology & Work*, *5*, 191-196.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree. *American Psychologist*, 64(6), 515.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. In: *Proceedings of the human factors society annual meeting* (Vol. 30, No. 6, pp. 576-580). Sage CA: Los Angeles, CA: Sage Publications.

Klein, G., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, *4*(3), 186–209.

Klein, K. J., Ziegert, J. C., Knight, A. P., & Xiao, Y. (2006). Dynamic delegation: Shared, hierarchical, and deindividualized leadership in extreme action teams. *Administrative Science Quarterly*, *51*(4), 590-621.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (2017). Situation awareness in team performance: Implications for measurement and training. *Situational awareness*, 63-76.

Schakel, J. K., & Wolbers, J. (2021). To the edge and beyond: How fast- response organizations adapt in rapidly changing crisis situations. *Human Relations*, *74*(3), 405–436.

Shu, Y., & Furuta, K. (2005). An inference method of team situation awareness based on mutual awareness. *Cognition, Technology & Work*, *7*, 272-287.

Stanton, N. A., Salmon, P. M., Walker, G. H., Salas, E., & Hancock, P. A. (2017). State-of-science: situation awareness in individuals, teams and systems. *Ergonomics*, *60*(4), 449-466.

Nja, O., & Rake, E. L. (2009). A discussion of decision making applied in incident command. *International Journal of Emergency Management*, *6*(1), 55-72.

Omodei, M. M., & Wearing, A. J. (1995). The Fire Chief microworld generating program: An illustration of computer-simulated microworlds as an experimental paradigm for studying complex decision-making behavior. *Behavior Research Methods, Instruments, & Computers*, 27(3), 303-316.

Polikarpus, S., Ley, T., Hazebroek, H., Edgar, G. K., Sallis, G., Baker, S., & Masip, A. F. (2022). Authoring Virtual Simulations to Measure Situation Awareness and Understanding. In *Proceedings of the 19th ISCRAM Conference – Tarbes, France May 2022 Hedi Karray, Antonio De Nicola, Nada Matta, Hemant Purohit*

Thoelen, F., Vastmans, J., Blom Andersen, N., Boehm, M., Holm, L., Arendtsen, B., ... & Walker, S. (2020). FireFront: A new tool to support training in Fireground Situation Awareness, Situation Understanding and

Bias. *International Fire Professional*, *34*, 34-39.

Thompson, J. D. (1967). *Organizations in action: Social science bases of administrative theory*. McGraw-Hill.

Treurniet, W., & Wolbers, J. (2021). Codifying a crisis: Progressing from information sharing to distributed decision-making. *Journal of Contingencies and Crisis Management*, *29*(1), 23-35.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, *185*(4157), 1124-1131.

Weick, K. E. (1993). The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly*, *38*(4), 628–652.

Weick, K. E., & Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly*, 357-381.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, 16(4), 409–421.

Wolbers, J., & Boersma, K. (2013). The common operational picture as collective sensemaking. *Journal of Contingencies and Crisis management*, 21(4), 186-199.

Wolbers, J., Boersma, K., & Groenewegen, P. (2018). Introducing a fragmentation perspective on coordination in crisis management. *Organization Studies*, *39*(11), 1521-1546.

Wolbers, J. (2022). Understanding distributed sensemaking in crisis management: The case of the Utrecht terrorist attack. *Journal of Contingencies and Crisis Management*, *30*(4), 401-411.

Collaborative Evaluation: Mapping Socio-Technical Patterns in an Emergency Exercise

Kjetil Rustenberg
Center for Integrated Emergency Management (CIEM), University of Agder, Norway
kjetil.rustenberg@uia.no

Kristine Steen-Tveit
Center for Integrated Emergency Management (CIEM), University of Agder, Norway
kristine.steen-tveit@uia.no

Bjørn Erik Munkvold
Center for Integrated Emergency Management (CIEM), University of Agder, Norway
bjorn.e.munkvold@uia.no

Jaziar Radianti
Center for Integrated Emergency Management (CIEM), University of Agder, Norway
jaziar.radianti@uia.no

ABSTRACT

Evaluating both exercises and real events can be challenging. This is especially evident in full-scale exercises, where a wide range of organizations must collaborate to solve a common problem. Information sharing, use of technology, and role understanding can be particularly challenging. In this article, we illustrate how academia can meaningfully engage in conducting and executing a comprehensive Emergency Exercise, ensuring that insights from prior research are evident from the outset. We employ a method where we co-create a Socio-Technical Network map together with practitioners. The method provides results that help visualize the complexity generated by large-scale collaboration, fostering involvement and engagement from both researchers and practitioners. Through this mapping, we gained valuable insights into barriers to Team Situation Awareness, technological possibilities and limitations, and a better understanding of various roles and their information needs.

Keywords

Full-scale exercises, collaborative evaluation, Team Situation Awareness, socio-technical network mapping

INTRODUCTION

Crises and unforeseen events are a part of reality. The frequency of these events can be related to contextual elements such as population density, climate change, and migration but also deliberate actions like terror and cyber-attacks (McCreight, 2023). A key feature is that these elements contain surprise (Deverell & Olsson, 2009). The wicked nature of a crisis challenges society to take relevant and adequate preparatory measures (Edzen, 2014; Owen et al., 2018), and handling crises demands various organizations to interact. Both defined emergency response organizations and organizations that do not usually handle crises must cooperate and form a new, temporary organization where cross-organizational collaboration is essential (Roud & Gausdal, 2019; Roud et al., 2021; Van Haperen, 2001). However, conducting exercises that prepare for collaboration in handling real events can be complex and challenging to achieve within the conditions that organizations typically operate under. There is an academic debate about the usefulness of full-scale exercises, which can be costly in terms of time and resources (Roud et al., 2021; Sørensen, 2017). Some have questioned the effectiveness of such exercises regarding individual learning outcomes (Berlin & Carlström, 2015; Borell & Eriksson, 2013). Previous studies also pinpoint the importance of differentiating between exercises focusing on individual skills and exercises focusing on collaboration and systems which prioritize the process to ensure effective collaboration (Roud et al., 2021). In the latter, elements such as information flow, communication channels, establishment of structures, and cross-organizational work beyond what one typically performs are central. A social-technical pattern within emergency

management describes the integrated systems of human behavior and technology that impact the response. It acknowledges that effective emergency management encompasses not solely technical components like technology, infrastructure, and protocols, but also social aspects such as individual situation awareness (SA) and Team Situation Awareness (TSA) (Endsley, 1995; Endsley, 2020). TSA is SA distributed across organizations, involving shared awareness of protocols, goals, and situational information (Sorensen & Stanton, 2011). Understanding patterns is essential for designing and implementing effective emergency management strategies that address both the technical and social aspects of the response.

Collaboration between academia and practice is essential for addressing these issues, but Davenport (1998) argues that this can be challenging. Academic research findings can be diffuse, or their points are already known by practitioners (Davenport, 1998). Furthermore, academia can be seen as an actor that tends to problematize rather than come up with solutions related to real-life problems. This means the practice fields may be reluctant to involve academia in their projects. This perspective can also be valid in exercise planning and execution.

This research aims to address the issue of effective collaboration and effective evaluation of full-scale exercises, something that requires a shared understanding of the social-technical patterns including both organizational-specific and shared information needs. We draw from experiences participating in a full-scale exercise that simulated a collision between two vessels in the waters off the coast of Southern Norway. This incident required a coordinated response effort from various emergency response organizations, municipalities, and volunteers. To approach these challenges this paper attempts to answer the Research Question (RQ): *How can the co-creation of Socio-Technical Network maps aid collaboration and understanding between researchers and practitioners during emergency exercises?* We used a Socio-Technical Network (STN) approach to initiate the work to uncover communication patterns and barriers to TSA. We mapped communication pathways, information technology in use, and plans and procedures in detail from the moment of the collision until the various organizations had initiated their work phase, in collaboration with practitioners as a part of the data collection. We argue that this approach is fruitful and provides a common understanding of the situation, making the evaluation work effectively. The data collection was conducted through participatory observations and qualitative, in-depth interviews with relevant stakeholders.

Considering the importance of TSA elements and Shared Mental Models (SMM), we seek to identify factors that strengthen collaboration and those that weaken it. SMM is further explained in the Theories section.

Identified elements of improvement from evaluation of the exercise will be part of the next planned full-scale exercise in which the authors will participate (Figure 1). Thus, we have the opportunity to follow up results and further validate our methodical approach.

BACKGROUND

The field of exercise methodology encompasses a wide range of approaches, and extensive literature exists discussing the efficacy of exercises in relation to real crisis management (Berlin & Carlström, 2015; Roud & Gausdal, 2019; Roud et al., 2021; Roud & Schmied, 2020; Sørensen, 2017). A common denominator among methodological approaches is that exercises aim to prepare the various actors to face real crises (Borell & Eriksson, 2013; McCreight, 2023). In crisis management, there are typically four or five stages or phases including preparedness, response, recovery, and mitigation. Some models also include prevention as a distinct phase (Coppola, 2006; McCreight, 2023). Prevention means actively taking measures to prevent crises and undesired events from occurring. Exercises are part of organizations' preparations for crises and undesired events, where exercises and training are included along with their plans and guidelines on how organizations can best ensure the best possible outcome. In Norway, different organizations have a high level of autonomy, which influences the amount of time and resources they dedicate to preparing for hypothetical scenarios. This includes scenarios based on risk analysis of past events. For instance, municipalities have varying degrees of available resources and personnel with specific roles related to exercises and preparations for undesired events. In larger municipalities, there are dedicated personnel with full-time positions related to crisis management. In comparison, crisis management may comprise only a limited role in smaller municipalities associated with personnel with completely different daily tasks. Also, within the established emergency response organizations, such as fire departments, police, and ambulance services, there is a high degree of autonomy. In addition, it is challenging to conduct satisfactory evaluations of emergency management exercises, as it also requires personnel to step out of their daily tasks, which can be costly and must be prioritized by the individual organizations (Steen-Tveit, 2020). Overall, this can challenge the opportunities for large-scale training.

In evaluation reports after major events in Norway, it is repeatedly pointed out that organizations struggle to learn from real events and exercises (Henriksen, 2022). Information sharing, TSA, and role understanding are recurring challenges. Despite identifying problem areas, the different organizations do not absorb the lessons, and it does not lead to real changes in behavior (Borell & Eriksson, 2013; DSB, 2020; Halvorsen, 2021; Henriksen, 2022;

Sønderland, 2012). In such cases, individuals or organizations continue to follow the same approach without critically examining or adjusting their methods, missing the opportunity for deeper learning and improvement that the events and exercises could provide (Deverell & Olsson, 2009; Roud et al., 2021). In the context of exercises, it is challenging to facilitate information sharing across organizational boundaries (Roud & Schmied, 2020). The established emergency response organizations have plans that ensure early sharing of important information so that each actor can have as accurate an understanding as possible of what they are responding to as early as possible. In Norway these are called Triple alert procedures (Steen-Tveit & Munkvold, 2021).

The procedure is anchored at the national level and has a historical basis in the unfortunate handling of a terrorist incident in 2013 (DSB, 2014). In this incident, the police knew that they were responding to an active knife stabbing while other emergency responders responded to what they believed was a bus accident. Procedures aimed at ensuring early similar situational understanding only apply to the collaboration between the fire department, ambulance service, and the police, as well as The Norwegian Joint Rescue Coordination Centres (NJRCC) and in some cases, the road traffic control center. In Norway, a collaborative model is used, where there are expectations that crises are handled by all relevant organizations in cooperation. Even if responsibility is not explicitly documented in procedures, the organization that normally has responsibility for a subject area or a geographical area should, as far as possible, assume this responsibility also in a crisis. This means that a cooperative response operation often becomes much larger than the established emergency response organizations, requiring planning and training for a sometimes very large, interdisciplinary, and cross-sectoral new organization.

THEORETICAL PERSPECTIVES

Team Situation Awareness (TSA) and Distributed Situation Awareness (DSA)

Theories on Situation Awareness (SA) are frequently used in emergency management, exercises, and evaluations following real-life events (DSB, 2020; Halvorsen, 2021). Shared situational understanding is essential when multiple actors interact and collaborate, but it can be challenging to achieve. Although reports have identified a lack of situational understanding as a challenge, barriers to obtaining shared situational awareness persist. Sharing relevant information and understanding the workflow processes within collaborating organizations are particularly challenging.

Endsley's 1995 SA model sheds light on SA as an individual, cognitive phenomenon and as a part of teamwork (Endsley, 1995). The theories for TSA refer to factors that promote shared situational understanding in collaborative teams. Central to successful collaboration is the sharing of information and familiarity among team members and building and maintaining SMM (Endsley, 2020; Endsley & Robertson, 2000). Mental models relate to internalized knowledge stored as schemas in long-term memory at an individual level. SMM relate to mental models being a product of collaboration and familiarity with the actors one collaborates with (Salas et al., 2015; Salas et al., 2005). With SMM, each team member can more easily understand the actions, needs, and future actions of other team members (Salas et al., 2005; Stout et al., 2017). A high degree of TSA is achieved when the team knows each other's needs and supports each other by sharing information and considering other members when performing an action (Endsley, 2020).

Stanton et al. recognize Endsley's theories regarding both individual and Team SA but emphasize the interaction between humans and technology. According to DSA (Distributed Situation Awareness) theory, SA is not primarily a cognitive phenomenon but an emerging phenomenon resulting from the interaction between both human and non-human actors in collaboration (Salmon, 2016; Stanton et al., 2006; Stanton et al., 2018). DSA can be characterized by the collective understanding of a situation among individuals or teams operating in a distributed environment. It involves sharing relevant information and knowledge to establish a common awareness of the situation. This shared awareness facilitates effective team coordination and allows synchronized actions and resource allocation across distributed entities. In DSA theory, it is argued that SA exists at a systems level and can be viewed as an emergent feature of collaborative systems (Salmon et al., 2017).

There is an ongoing debate about whether Stanton et al.'s perspectives contribute a new theoretical perspective or if this is part of Endsley's described Distribution of Situational Awareness. In this paper, we choose to utilize perspectives from both theories and believe that they complement each other at a higher level (Endsley, 2015; Salmon et al., 2008).

SCENARIO DESCRIPTION

Large-scale exercises aimed at testing collaboration at the local system level are a recurring event in Norway, occurring annually and in multiple locations. In the exercise used in this article, a maritime incident was simulated. The scenario involved a collision between two medium-sized vessels in the archipelago area between two

municipalities in southern Norway. One vessel had 71 persons on board, while the other had 31 persons, totaling 102 actors who played the roles of shipwreck victims. The exercise commenced at 09:18 AM when one of the vessels sent out a MAYDAY distress call, which was intercepted by the Coastal Radio. The message reported a collision, water ingress in one vessel causing it to list and risk sinking, and smoke and fire onboard the other vessel. The message was immediately relayed to the NJRCC, who initiated the notification of police, fire brigade, and healthcare services in the relevant municipalities. Notification of emergency responders geographically linked to the incident follows national Search and Rescue (SAR) standard operating procedures (SOP). These procedures have similar wording and layouts to ensure an early common situational understanding among emergency responders.

Local tactical teams were then notified to respond to the scene, and a common radio frequency was established for coordination. Different emergency response organizations have independent responsibilities to alert other necessary resources, as specified in written SOPs available to operators. For instance, the Emergency Medical Communication Center (EMCC) is responsible for alerting ambulances and doctors in municipal duty rosters and keeping nearby hospitals informed. The police are responsible for leading local operations and also for notifying municipalities, volunteer organizations, and others who may have responsibilities related to handling the incident.

In the early stages of the operation, it was decided that the vessels would prioritize going to two different ports in each municipality. The land-based aspect of the incident had to be coordinated as both the same and as two separate incidents. Two centers for evacuees and relatives were established, and there was a need for information from local actors, such as the press, at both locations. However, only one of the vessels made it to port. For the other vessel, it was simulated that it sank, and passengers came ashore on various islets and rocks.

Overall, the scenario ensured that a relevant selection of actors that one would normally expect in a real scenario was notified and had the opportunity to practice cooperation and communication. There was also relevant stress on the technical structures, including intensive use of the Norwegian Critical Communication Network (NCCN), which was pushed close to its maximum capacity.

The exercise concluded at 1:00 PM. Individual organizations sporadically evaluated it through Hot Wash sessions, but none took the initiative for a larger, cross-organizational evaluation. We, as researchers, distributed a Survey and gathered input from observers.

METHODOLOGY AND DATA COLLECTION

Prior to the exercise, as researchers, we were engaged by the exercise committee to work with them to develop clear exercise objectives that could be followed throughout the exercise and evaluated afterwards. Figure 1 shows how we were engaged throughout the exercise.

The exercise objectives were the result of a workshop held beforehand, as well as a tabletop exercise focusing on identifying whether each organization's plans and procedures aligned with those of the other collaborating organizations. The exercise objectives are shown in Table 1.

Table 1. Exercise objectives

Goal #	Description
Overall goal 1:	Conduct the exercise with unified contingency plans, wherein participating organizations collaborate towards common objectives and optimal resource utilization across agencies.
Overall goal 2:	Ensure that the rescue service always has the necessary information to create a team situational awareness.
Overall goal 3:	Ensure coordinated media and information services at all levels to provide correct information at the right time.
Overall goal 4:	Optimize resource utilization by leading, coordinating, and supporting rescue resources at strategic, operational, and tactical levels.

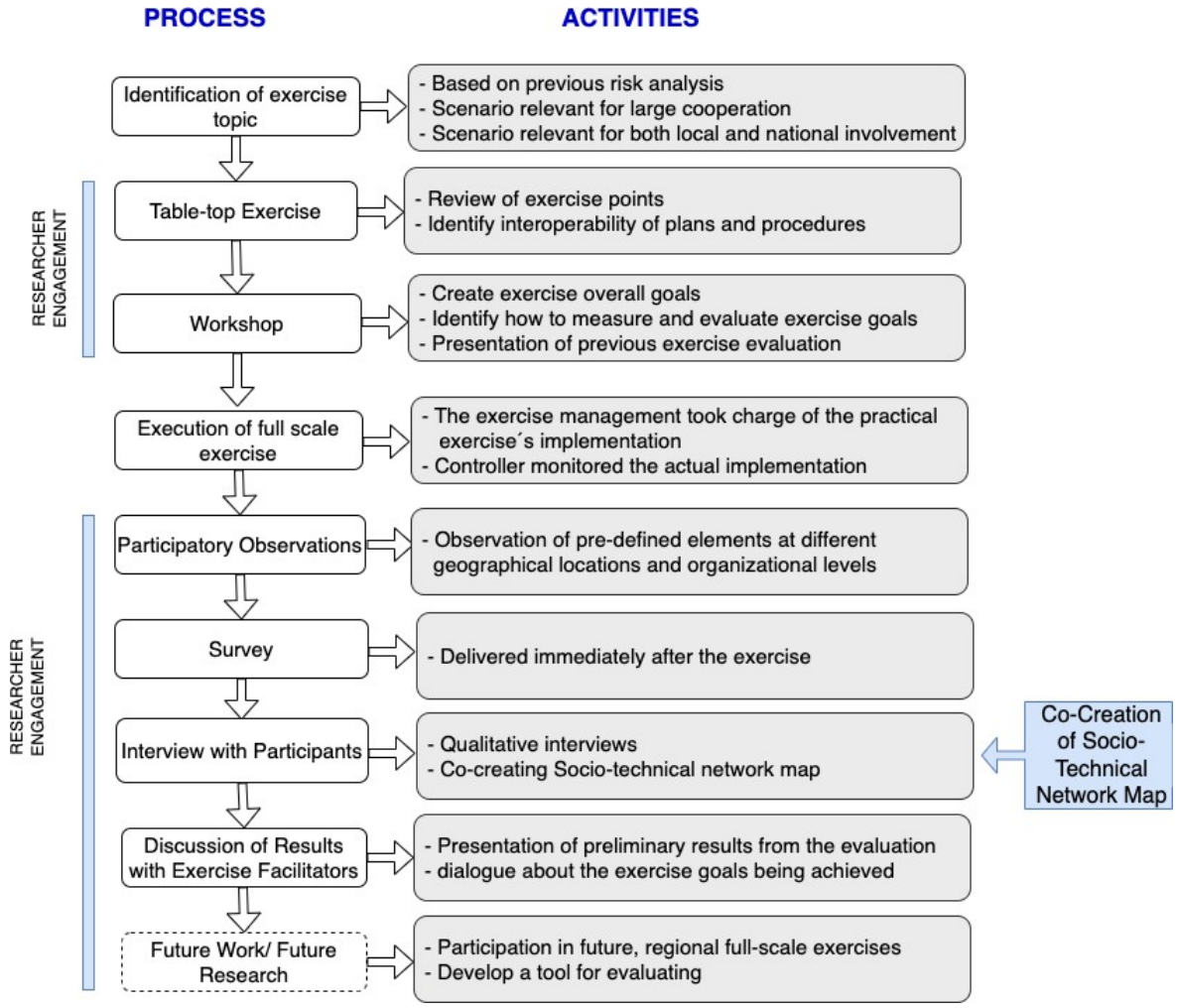


Figure 1. Researcher engagement in exercise

Data from the exercise was collected in four different ways: 1) A total of seven individuals, including three of the authors of this study observed the exercise at various geographical locations. At predefined times, observations related to the exercise objectives were recorded, using an observation guide to direct attention towards predefined elements and minimize observer bias; 2) Following the exercise, a survey was distributed to approximately 200 respondents, of whom 154 responded. The survey contained both closed and open-ended questions; 3) Selected findings from both observations and the survey are now being followed up with qualitative interviews. At the time of writing this article, a total of 13 interviews have been conducted. The interviews follow an interview guide aimed at clarifying respondents' role understanding, use of information systems, information sharing and communication, and experiences with evaluation and learning from exercises and real events. The length of each interview varies from 50 minutes to 150 minutes; 4) Before the interviews, we mapped the notification sequences, information systems in use, communication channels, and SOPs using MindManager[1]. This commercial mind-mapping software allows users to visualize information in mind maps and flowcharts. Together with the respondents, we enriched the initial mapping and co-created a network map that visualizes how information flows and how it is supported or not supported by adequate technical solutions and SOPs, thus supporting or creating barriers to TSA. Figure 2 shows the initial mapping of notification sequence, and Figure 3 shows initial notification with communications relations to VHF Maritime channels and the NCCN. Note that the original mapping is done in Norwegian. In this article, the figures demonstrate our methodical approach, but details have been removed for security and confidentiality reasons.

[1] https://www.mindmanager.com/en/

Socio-Technical Network (STN) Mapping

To approach an evaluation of whether the overall goals were achieved (Table 1), we, together with the respondents, have conducted an STN mapping as a part of our qualitative interviews (Figure 1). The mapping is delimited in time from the exercise's inception until the notification phase was over and the various organizations began the work phase. To gain an overview and to map information patterns and identify how information technology supported the human actors' SA, our mapping aims to identify both the human and the technical, or non-human, actors. Our approach is inspired by Salmon et al.'s (2017) approach to representing DSA, and also the Delphi method (Linstone & Turoff, 1975) where the "informed judgements" and expert opinions from practitioners are highly valued (Linstone & Turoff, 1975; Plotnick et al., 2011). The use of Social-technical network analysis in connection with emergency management is not uncommon. As an example, Houghton et al. (2006) seeks to elucidate and categorize archetypal networks in Command&Control networks. It suggests that social network analyses can be a valuable part of general studies of command and control. Mapping social networks along with other elements of an information system seems timely. In this article, we consciously call our approach a STN-mapping to clearly emphasize a connection between human and non-human actors, procedures, and communication pathways, and to clarify a distinction between Mapping and Analysis. Our mapping is a part of our analysis, but not an analysis on its own. In social networks, the relationship and/or flow of communication between human actors as individuals and groups and non-human actors such as computers and other data processing devices are plotted as connections between entities (Houghton et al., 2006).

Our STN mapping serves as a valuable tool in an interview situation. Collaboration between practitioners and researchers is essential, and input from practitioners can immediately be translated into a product recognizable to the practitioner and usable by the researcher to clarify what actually transpired. Thus, both parties have ownership of the end product. As a premise, only respondents in the interview were allowed to comment on the parts of the mapping they were involved in. Any objections or comments of others were noted and discussed with the relevant parties. The idea of the researcher being fully objective is an obsolete idea, at least from an interpretivist understanding of epistemology (Klein & Myers, 1999; Wynn Jr & Williams, 2012). An interview situation involving a researcher and a respondent represents an asymmetric power relationship, where the researcher holds a position where (s)he both leads the dialogue and interprets the information that emerges. Our method alleviates this asymmetry and the product we co-create functions as a boundary object (Star, 1989).

PRELIMINARY RESULTS AND DISCUSSION

In our analysis, we observe that the initial resource notification worked well in the very first sequence (Figure 2). In this part of the notification process, a SOP regularly drilled by established emergency response organizations was used. The SOP is well-known among all personnel and is rooted in national guidelines. However, as the complexity of the incident increased, the established emergency response organizations encountered increasing challenges in continuing resource notification. In cases where automated responses managed the notification, the process itself was unproblematic. However, when operators had to manually notify using, for example, call lists, there were failures in the notification process. For instance, the EMCC conducted macro alerts via the emergency network, where predefined NCCN macros alerted all healthcare resources considered relevant within the municipal boundaries in a cascading operation. This included hospitals closest to the incident site. Operators could verify that the alert had been sent and the call received without verbal confirmation over the communication system. However, organizations responsible for notifying volunteer organizations depended on operators physically calling each organization, providing relevant information, confirming receipt of the mission, and their deployment. In the exercise, we observed that this process did not function as intended. Some operators reported that they were overwhelmed with incoming calls, and that their organization's internal information needed in some cases forced personnel to prioritize this over following protocol. The STN network mapping clearly highlighted the causes of certain functions becoming bottlenecks in the flow of information. Here, we identified issues that challenged the overall exercise goals 1 and 2 (Table 1), such as organizations that were not present on the NCCN and, therefore, omitted from the information loop (Figure 3).

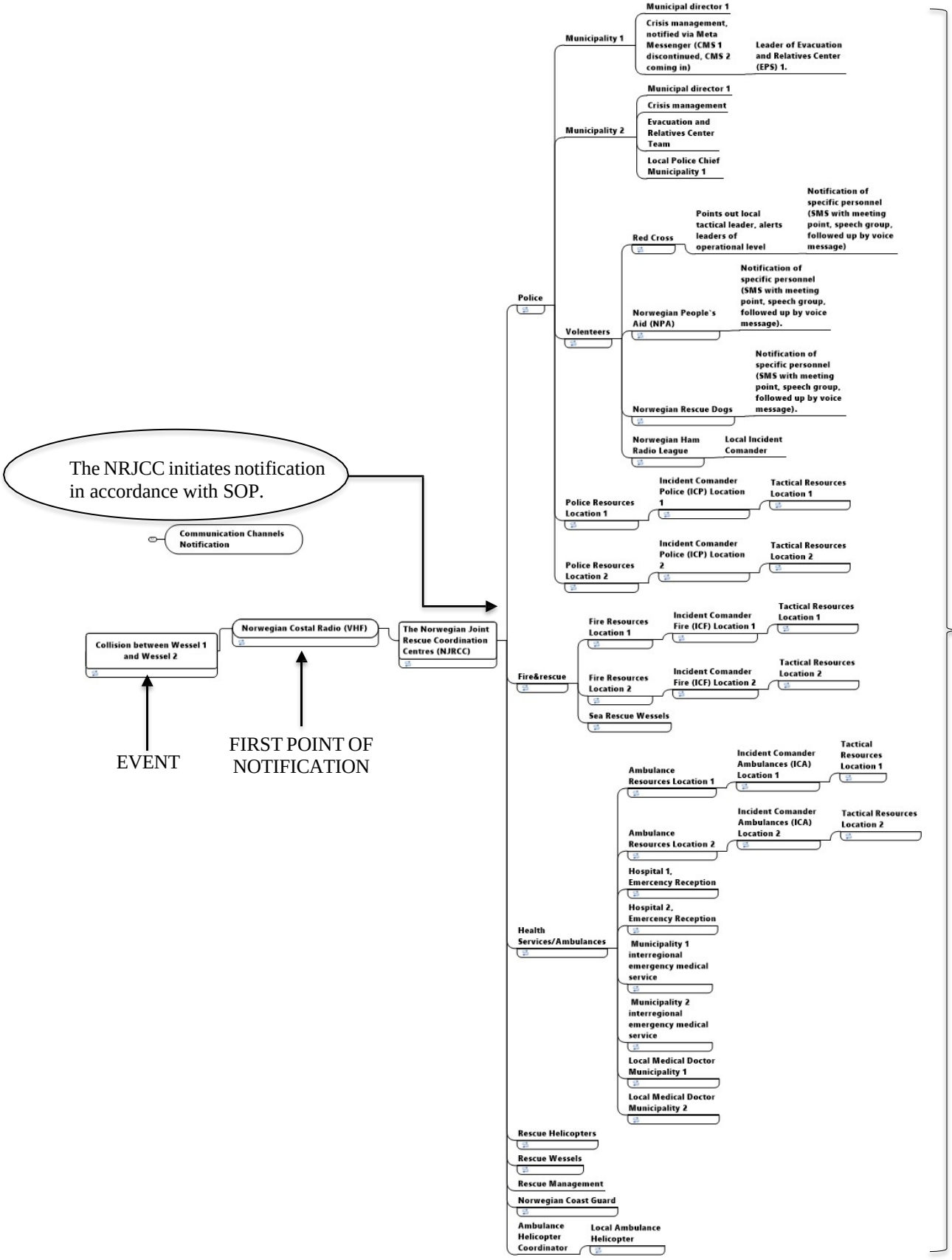


Figure 2. Socio-Technical Network (STN) Mapping of alerting phase without communications relations

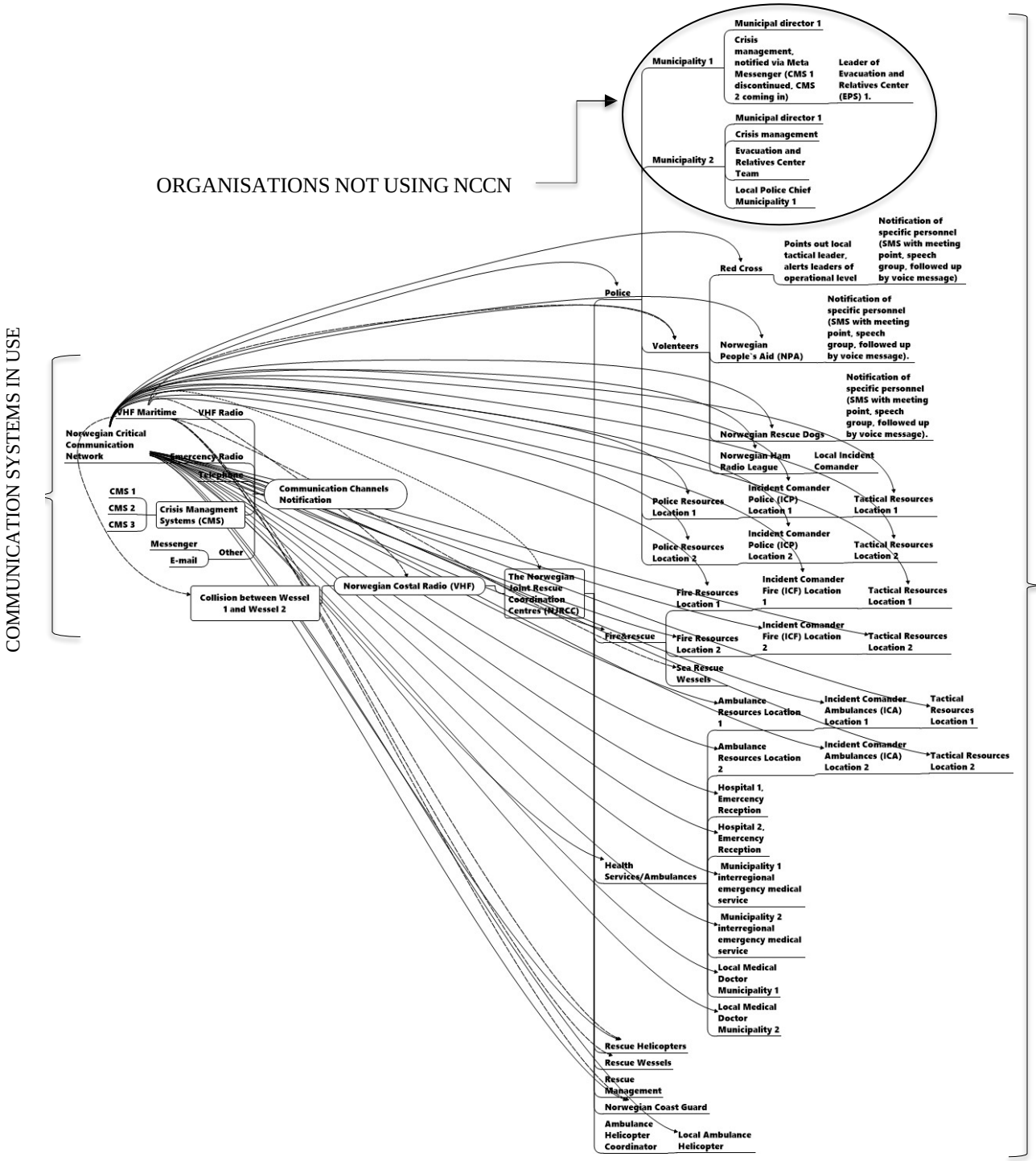


Figure 3. Socio-Technical Network (STN) Mapping with communications relations

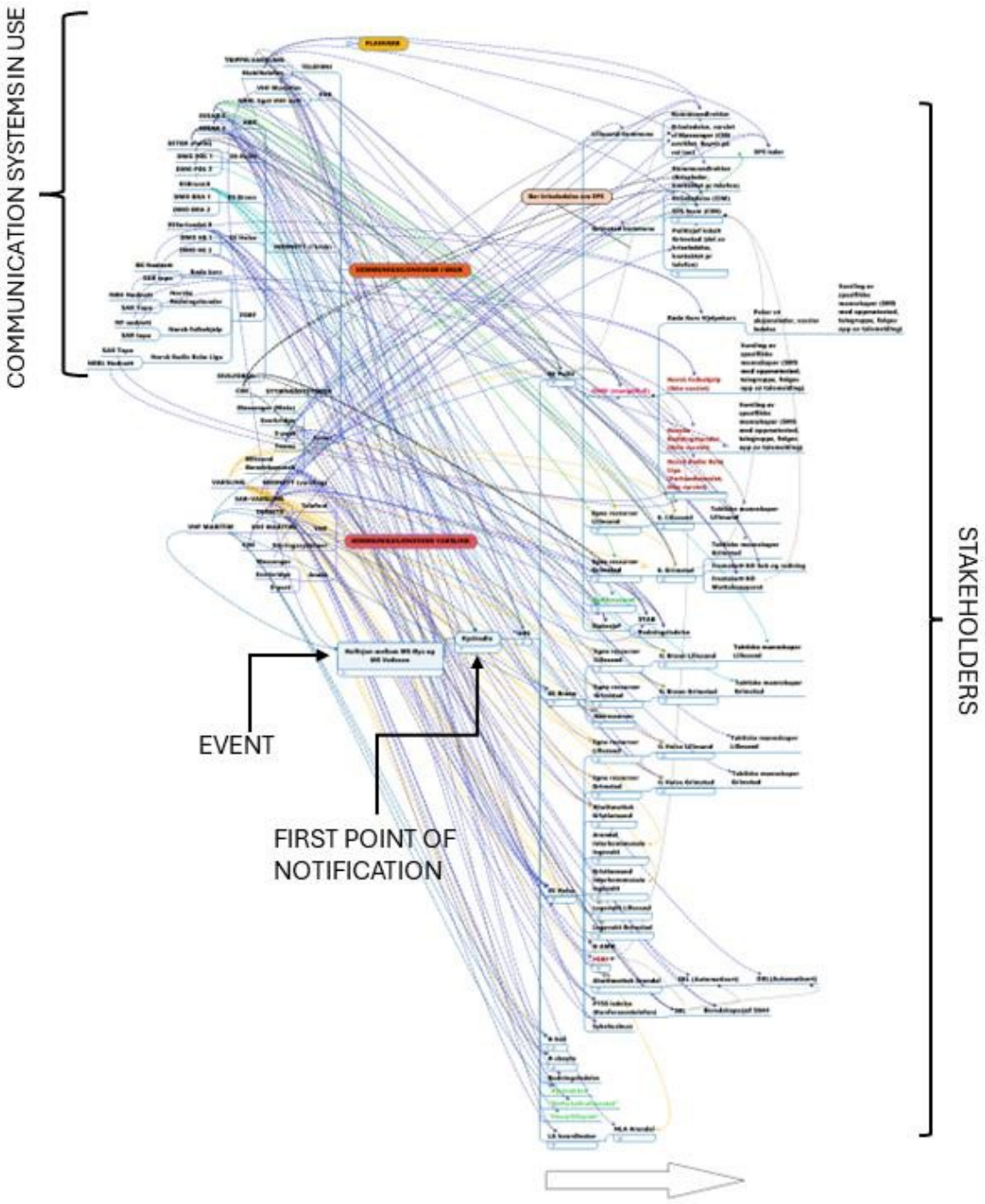


Figure 4. Socio-Technical Network (STN) Mapping full-blown from actual exercise

The mapping effectively illustrates the diversity of information systems operational during the exercise and to what extent they supported or hindered the achievement of a high level of TSA within the collaboration (Figure 4). In the initial alerting phase, eight different communication systems were in use, none of them overlapped. Consequently, interoperability posed a challenge, particularly affecting organizations unfamiliar with emergency network radios. This also challenged overall exercise goal 2 since it hampered the rescue service in getting necessary information to create TSA (Table 1).

An important function of the NCCN is the ability to set up different call groups or "digital rooms." It is possible to set up several call groups, both agency-specific and cross-organizational, during one emergency operation (Tveit et al., 2022). We observed that the NCCN structure was challenged as the number of call groups increased as the exercise progressed. Initially, agencies were notified by phone, and tactical teams were alerted in their own agency-specific call groups. They were then moved to a common call group in the very first part of the response phase. When the first resources arrived at the scenes, there were six active call groups across multiple base stations. As the operational phase progressed, this number grew to twelve call groups, with only one base station.

Additionally, several tactical teams communicated directly using direct mode communication, (i.e., communication from terminal to terminal, bypassing a common base station which releases network capacity). This fragmentation of communication pathways and information channels meant no single common location where all involved actors could expect to receive important information relevant to the collaboration.

The mapping also revealed deficiencies in the procedures used by the actors, demonstrating how, in some cases, procedures had to be abandoned because they were not practically applicable in the real world. As such, the mapping serves to operationalize procedures further by identifying why they did not function satisfactorily. For example, the procedure supporting municipal employees in establishing centers for evacuees and relatives required registration of evacuees in two ways, both in an Excel spreadsheet and in a separate system. However, the separate system had usability issues, such as the end-user receiving no information or confirmation that entered personal data was sent and received by the appropriate party. Additionally, evacuees could enter their information into this system by scanning an available QR code. This led to people being double-registered or not registered at all, without those on the front lines having oversight. The solution was to abandon the entire system and enter patient data into an Excel spreadsheet, which was then emailed to the central police. Another example was that, at one point, there were challenges associated with disseminating information from the centers for evacuees and relatives to other relevant stakeholders. Initially, there were no apparent reasons for this beyond the high workload of the centre's personnel. As completing the STN-map progressed, the issue became visually and explicitly apparent, leading to further exploration and contact with the center's leader. In dialogue with them, it emerged that they were receiving conflicting instructions about where they should be located, and at one point, the leader was ordered by the on-site police to relocate to the reception site where the involved vessel was arriving. With that, the contact point other organizations utilized to reach the center also disappeared, compromising the flow of information.

The collaboration process with practitioners to develop a STN map provides a thorough and objective illumination of potential barriers to Team Situational Awareness. In the co-creation process, we highlighted the complexity that extensive collaboration entails, including diverse roles, inter-organizational work across sectors, and various communication technologies with varying degrees of interoperability. The mapping also identified organizations that fell outside the majority, thus risking missing crucial information and consequently having a lower level of TSA.

We have presented our findings, illuminated through the mapping, to exercise leaders and leaders from the involved organizations. The feedback has been positive, and there is a plan for greater involvement in the next exercise to be implemented.

Both data collection and analysis processes in this type of study are demanding. As two researchers have a practitioner background, we quickly gained access to organizations where insight is typically hard to obtain. An example is that two of the researchers have a national personal security clearance that includes exposure to classified and sensitive information. This placed us in a dual role where, on the one hand, we gained rapid access to unrestricted information, while on the other hand, we individually and personally had to take responsibility for not exposing data that should not be disclosed. The solution to this is to collaborate closely with leaders from the various organizations, ensuring redundancy in the control of the information exposed for scientific purposes.

CONCLUSION AND FUTURE WORK

If academia is to be involved in such exercises as we describe, we find it highly beneficial to have early involvement. By being able to shape the wording of the exercise objectives and anchor them in the real problems experienced by the collaboration, we found that the entire process of ensuring that the various actors were involved in the exercise with the same focus, as well as the evaluation work, became much easier. We see the value of having a presence in the field of practice and approaching experts from various disciplines with a certain humility. The interaction between academia and practice can be highly fruitful and valuable. Still, it can easily be compromised if there are questions about ulterior motives or if trust between researchers and practitioners is undermined. Using a simple STN mapping, as demonstrated in this article, provides an element of objectivity. The map serves as a boundary object between practitioners with experience from exercises and real events, and us as researchers. The collaboration in recreating what actually happened, how it happened, and what support each actor received from other actors, technology, or predefined procedures allows for a good and candid dialogue between practice and academia, which we argue is valuable. The network mapping effectively visualized the complexity of some of the roles that needed to be fulfilled (Figure 4). In some cases, this visualization led to a demystification of acknowledging mistakes. Some actors were surprised to see how they were part of such a large interaction with other actors and technology. The findings can be used to refine the use of technology, with the understanding that TSA is necessary for the cooperation to achieve its goals when working together in a collaborative, large, fragmented, and temporary organization. We expect that a visualization similar to the one we

have conducted will strengthen and develop the SMM of the involved actors. Preliminary feedback suggests that the co-creation of the STN map and early researcher involvement have already contributed to a greater understanding of participants' roles, the information needs of collaborating organizations, and technological possibilities and limitations.

In our future research we will further develop and test a tool for evaluating post-exercise measures. We will also seek to identify Information Trading Zones (ITZ) within the various communication structures, and thus further validate a preliminary framework introduced in (Rustenberg et al., 2024). ITZ in the context of Emergency Management can be defined as a space where various actors "trade" their ideas about the relevance of available crisis information (Boersma et al., 2014; Boersma et al., 2012). Central to this is identifying barriers to the involved stakeholders' and organizations' ability to develop and maintain TSA. The STN mapping described in this paper will be further developed to visualize how the information trading during the exercise actually occurred and the limitations and potential of the Information Systems in use.

REFERENCES

Berlin, J. M., & Carlström, E. D. (2015). Learning and Usefulness of Collaboration Exercises: A Study of The Three Level Collaboration (3LC) Exercises Between The Police, Ambulance, and Rescue Services. *International Journal of Mass Emergencies & Disasters*, *33*(3).

Boersma, F., Comfort, L., Groenendaal, J., & Wolbers, J. (2014). Incident command systems: A dynamic tension among goals, rules and practice. *Journal of Contingencies and Crisis Management*, *22*(1), 1-4.

Boersma, K., Wagenaar, P., & Wolbers, J. (2012). Negotiating the'Trading Zone'. Creating a Shared Information Infrastructure in the Dutch Public Safety Sector. *Journal of Homeland Security and Emergency Management*, *9*(2).

Borell, J., & Eriksson, K. (2013). Learning effectiveness of discussion-based crisis management exercises. *International Journal of Disaster Risk Reduction*, *5*, 28-37.

Coppola, D. (2006). *Introduction to international disaster management*. Elsevier.

Davenport, T. H. (1998). Putting the enterprise into the enterprise system. *Harvard business review*, *76*(4).

Deverell, E., & Olsson, E.-K. (2009). Learning from crisis: A framework of management, learning and implementation in response to crises. *Journal of Homeland Security and Emergency Management*, *6*(1).

DSB. (2014). The Valdress Express: Evaluation of the authorities' handling of the November 2013 incident. The Norwegian Directorate for Civil Protection.

DSB. (2020). *Evaluation of The Viking Sky incident*. The Norwegian Directorate for Civil Protection. https://www.dsb.no/globalassets/dokumenter/rapporter/evaluering_av_viking_sky_hendelsen.pdf

Edzen, S. (2014). Table-top exercises for emergency management: tame solutions for wicked problems. 2014 47th Hawaii International Conference on System Sciences,

Endsley, M. R. (1995). Toward a theory of situational awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32-64. https://doi.org/10.1518/001872095779049543

Endsley, M. R. (2015). Situational Awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 4-32.

Endsley, M. R. (2020). Situation Awareness in Teams: Models and Measures. In *Contemporary Research* (pp. 1-27). CRC Press.

Endsley, M. R., & Robertson, M. M. (2000). Situation awareness in aircraft maintenance teams. *International Journal of Industrial Ergonomics*, *26*(2), 301-325.

Halvorsen, J. (2021). *EVALUATION: The rescue operation and the acute handeling of the crisis during the quick clay landslide at Gjerdrum*.

Henriksen, S. V. B., Torgeir; Hoel, Linda; Kruke, Bjørn Ivar; Myhrer, Tor-Geir; . (2022). *Evaluation of PST's and the police's handling in connection with the Kongsberg incident on 13th October 2021*.

Houghton, R. J., Baber, C., McMaster, R., Stanton, N. A., Salmon, P., Stewart, R., & Walker, G. (2006). Command and control in emergency services operations: a social network analysis. *Ergonomics*, *49*(12-13), 1204-1225.

Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS quarterly*, 67-93.

Linstone, H. A., & Turoff, M. (1975). *The delphi method*. Addison-Wesley Reading, MA.

McCreight, R. (2023). *An introduction to emergency exercise design and evaluation*. Rowman & Littlefield.

Owen, C., Brooks, B., Curnin, S., & Bearman, C. (2018). Enhancing learning in emergency services organisational work. *Australian Journal of Public Administration*, *77*(4), 715-728.

Plotnick, L., Turoff, M., & White, C. (2011). Partially distributed emergency teams: considerations of decision support for virtual communities of practice. *Supporting Real Time Decision-Making: The Role of Context in Decision Support on the Move*, 203-220.

Roud, E., & Gausdal, A. H. (2019). Trust and emergency management: Experiences from the Arctic Sea region. *Journal of Trust Research*, *9*(2), 203-225.

Roud, E., Gausdal, A. H., Asgary, A., & Carlström, E. (2021). Outcome of collaborative emergency exercises: Differences between full-scale and tabletop exercises. *Journal of Contingencies and Crisis*

Management, *29*(2), 170-184.

Roud, E., & Schmied, J. (2020). Emergency Collaboration Exercises and Learning: Experiences from the Arctic. In *Crisis and Emergency Management in the Arctic* (pp. 180-211). Routledge.

Rustenberg, K., Radianti, J., & Gjøsæter, T. (2024). Adopting a Trading Zone Framework to Emergency Management. In Proceedings of the 57th Hawaii International Conference on System Sciences. HICSS.

Salas, E., Shuffler, M. L., Thayer, A. L., Bedwell, W. L., & Lazzara, E. H. (2015). Understanding and improving teamwork in organizations: A scientifically based practical guide. *Human resource management*, *54*(4), 599-622.

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "big five" in teamwork? *Small group research*, *36*(5), 555-599.

Salmon, P. M. (2016). *Distributed situational awareness : theory, measurement and application to teamwork.* Routledge.

Salmon, P. M., Stanton, N. A., & Jenkins, D. P. (2017). Distributed situation awareness: Theory, measurement and application to teamwork. CRC Press

Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., Jenkins, D. P., McMaster, R., & Young, M. S. (2008). What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, *9*(4), 297-323.

Sønderland, O. D., Bjørn Kåre; Eriksen, Egil; Frøyland, Espen; Krogstad, Terje; Paulsen, Øystein Stavdal;. (2012). *22 July 2011: Evaluation of the police's efforts*.

Sørensen, J. L. (2017). *Norwegian maritime crisis collaboration exercises: are they useful?* Northcentral University.

Sorensen, L. J., & Stanton, N. A. (2011). Is SA shared or distributed in team work? An exploratory study in an intelligence analysis task [Article]. *International Journal of Industrial Ergonomics*, *41*(6), 677-687. https://doi.org/10.1016/j.ergon.2011.08.001

Stanton, N. A., Stewart, R., Harris, D., Houghton, R. J., Baber, C., McMaster, R., Salmon, P., Hoyle, G., Walker, G., & Young, M. S. (2006). Distributed situational awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics*, *49*(12-13), 1288-1311.

Stanton, N. A. D., Salmon, P. D., & Walker, G. H. D. (2018). *Systems thinking in practice: applications of the event analysis of systemic teamwork method*. CRC Press.

Star, S. L. (1989). The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In *Distributed artificial intelligence* (pp. 37-54). Elsevier.

Steen-Tveit, K. (2020). Identifying information requirements for improving the common operational picture in multi-agency operations. Proceedings of the 17th ISCRAM Conference,

Steen-Tveit, K., & Munkvold, B. E. (2021). From common operational picture to common situational understanding: An analysis based on practitioner perspectives. *Safety science*, *142*, 105381.

Stout, R. J., Cannon-Bowers, J. A., & Salas, E. (2017). The role of shared mental models in developing team situational awareness: Implications for training. In *Situational awareness* (pp. 287-318). Routledge.

Tveit, K. S., Munkvold, B. E., & Radianti, J. (2022). Using live video for communication between lay bystanders and emergency dispatchers in command and control centres. *International Journal of Emergency Management*, *17*(2), 154-176. https://doi.org/10.1504/IJEM.2021.122930

Van Haperen, K. (2001). The value of simulation exercises for emergency management in the United Kingdom. *Risk Management*, *3*, 35-50.

Wynn Jr, D., & Williams, C. K. (2012). Principles for conducting critical realist case study research in information systems. *MIS quarterly*, 787-810.



ISCRAM 2024
MÜNSTER, GERMANY

Geographic Information Science (GIS) for Crisis Management

Track Chairs

Michael A. Erskine, MTSU

Joy Garfield, Aston University

Diego Pajarito Grajales, University of Glasgow

Andrés Díaz López, Arizona State University Olaf

Steenbergen, Netherlands Red Cross

QuakeSet: A Dataset and Low-Resource Models to Monitor Earthquakes through Sentinel-1

Daniele Rege Cambrin*
Politecnico di Torino
daniele.regecambrin@polito.it

Paolo Garza
Politecnico di Torino
paolo.garza@polito.it

ABSTRACT

Earthquake monitoring is necessary to promptly identify the affected areas, the severity of the events, and, finally, to estimate damages and plan the actions needed for the restoration process. The use of seismic stations to monitor the strength and origin of earthquakes is limited when dealing with remote areas (we cannot have global capillary coverage). Identification and analysis of all affected areas is mandatory to support areas not monitored by traditional stations. Using social media images in crisis management has proven effective in various situations. However, they are still limited by the possibility of using communication infrastructures in case of an earthquake and by the presence of people in the area. Moreover, social media images and messages cannot be used to estimate the actual severity of earthquakes and their characteristics effectively. The employment of satellites to monitor changes around the globe grants the possibility of exploiting instrumentation that is not limited by the visible spectrum, the presence of land infrastructures, and people in the affected areas. In this work, we propose a new dataset composed of images taken from Sentinel-1 and a new series of tasks to help monitor earthquakes from a new detailed view. Coupled with the data, we provide a series of traditional machine learning and deep learning models as baselines to assess the effectiveness of ML-based models in earthquake analysis.

Keywords

Remote Sensing, Earthquake Monitoring, Deep Learning, Machine Learning, Regression, Detection, Change Detection

INTRODUCTION

The application of modern machine learning and deep learning solutions to remote sensing is strictly linked to data availability. Thanks to the launch of Landsat-8 (Wulder et al., 2019), MODIS (Justice et al., 2002), and Sentinel missions (Main-Knorn et al., 2017; Torres et al., 2012), general-purpose datasets were released in recent years (Helber et al., 2019; Sumbul et al., 2019; Y. Wang et al., 2022). Emergency management can benefit from these solutions. However, only a few datasets of satellite imagery were publicly released for a limited number of emergency management tasks (e.g., flood detection (Bonafilia et al., 2020) and burned areas delineation (Cambrin et al., 2023)). Ad-hoc image datasets are unavailable for many emergency management tasks. Among the others, public datasets of satellite imagery tailored to training machine learning models for earthquake analysis have not been released to our knowledge.

Earthquake monitoring is mainly based on seismometers that actively record seismic waves to detect, understand the strength, and identify the origin of these events. However, we cannot have global capillary coverage using this instrumentation. Conversely, satellite instruments have a broader view over the globe. Hence, they can help reach remote places worldwide and understand which areas were affected by the earthquakes without sending people to assess the damages and without using seismometers. Tasks such as identifying affected areas and estimating magnitude can be solved offline by collecting satellite images and then applying machine learning models in a data center. Alternatively, the inference can be made directly onboard the satellites using low-resource models, avoiding storing large amounts of data and enabling the possibility of transmitting only a few significant information (e.g.,

*corresponding author

coordinates of the damaged area, the estimated magnitude, etc). This raises another crucial challenge linked to the resource availability of the used systems: employing a large model requires computational resources unavailable on board. Creating an accurate model is not the only requirement. Low resource consumption is mandatory, and memory, CPU, and energy must be used effectively.

In this context, we propose a new dataset of more than a hundred earthquakes around the Earth, containing thousands of tri-temporal samples of the same event using Sentinel-1 data and ISC annotations. Moreover, we assess the effectiveness of shallow and deep learning models to solve tasks related to earthquake monitoring.

Our contributions can be summarized as follows:

- We propose a series of tasks to analyze earthquake-affected areas that can be solved by applying machine learning solutions to Sentinel-1 data.
- We evaluate a series of deep learning and shallow machine learning solutions in a subset of the proposed tasks.
- We pose the tasks as a balance between performance and resource consumption to democratize the proposed solutions and make feasible the application to low-resource devices.
- We publicly release a collection composed of tri-temporal time-series images, allowing the possibility to analyze time-related changes.

For reproducibility, we released the code for the experiments at https://anonymous.4open.science/r/quakeset-4D73/. The dataset is available at https://zenodo.org/records/10423411 to allow the community to experiment with it. Both code and data are released anonymously to comply with the double-blind review process.

RELATED WORKS

Image Processing through Deep Learning

The rise of convolutional neural networks and, more recently, vision transformers significantly improved the capability of processing images. The great data availability and the use of pre-trained neural networks posed deep learning models as state-of-the-art solutions in computer vision over classical machine learning models in benchmark datasets like ImageNet (Deng et al., 2009). The increasing performance of these models is often linked to the increased computational costs at training time and in inference (Xie et al., 2021). Dealing with real-time constraints or low-resource devices poses the challenge of balancing accuracy with resource consumption. Developing specific deep learning models like MobileNets (Howard et al., 2017; Sandler et al., 2018) or MobileViTs Mehta and Rastegari, 2021, 2022 were strictly linked to these constraints. Other solutions were designed in different fashions, including the possibility of employing less accurate real-time versions or the heavier and best-performing ones (Z. Liu et al., 2022; W. Wang et al., 2021; Woo et al., 2023; Xie et al., 2021). Finally, the effectiveness of the deep learning models in computer vision is highly impacted by data availability. Without ad hoc collections of images, the results decrease rapidly when considering domains for which tailored datasets are unavailable or limited.

Image Processing and Deep Learning in Crisis Management

Using textual information to analyze the evolution of ongoing emergencies has proven effective with the solutions proposed in the Incident Stream (McCreadie et al., 2019) and Crisis Facts (McCreadie et al., 2023) tracks of TREC. Although, also the analysis of images taken from social media has proven to be effective in the identification of areas affected by fires (Daly & Thom, 2016), earthquakes, hurricanes, and typhoons (Alam et al., 2017; Nguyen et al., 2017). The training of deep convolutional neural networks grants improved results over traditional techniques like Bag-Of-Visual-Words (Nguyen et al., 2017). The release of Incidents1M (Weber et al., 2022) with around a million images of 43 incidents granted the possibility to pre-train vision transformers and finetune existing pre-trained ones (Long et al., 2023). The significant drawback of all existing solutions is that they rely on social media and RGB images. RGB images can limit the comprehension of certain aspects of the events due to the limitations of the visible spectrum. For instance, using the infrared spectrum in Sentinel-2 helps monitor green areas (Drusch et al., 2012), being vegetation more sensitive to these frequencies. Moreover, using social media data also excludes the possibility of reaching sites with limited human presence or low-quality or damaged communications since they rely on information people send over social media. For these reasons, we propose to use satellite data, specifically Sentinel-1 images.

Machine Learning in Remote Sensing

The advancement of neural networks greatly benefits applications in the remote sensing field (Yuan et al., 2020; X. X. Zhu et al., 2017). The networks designed for RGB images have proven adaptable to other spectral bands like infrared and ultra blue. An extensive series of classical computer vision tasks, such as image classification and segmentation, were adapted to the remote sensing domain. The BigEarthNet (Sumbul et al., 2019) and EuroSAT (Helber et al., 2019) datasets pose the problem of landcover detection as an image classification one. The Seaships dataset (Shao et al., 2018) analyzes an object detection problem, while CaBuAr (Cambrin et al., 2023) analyzes a semantic segmentation problem in the burned area identification context. In addition to the extra spectral bands per image, in many cases, these datasets contain spatiotemporal information granting the possibility of solving tasks like change detection, which aims to find in supervised or unsupervised way the difference between two samples of the same spatial area taken at two different timestamps (Asokan & Anitha, 2019). The WHU dataset (Ji et al., 2018) deals with this challenge for urban planning.

Machine Learning in Seismology

There is an increasing interest in applying machine learning and deep learning in the seismology field (Chandrakumar et al., 2023). The current machine-learning solutions were mainly applied to earthquake seismic waves (P-waves and S-waves) to detect an earthquake in real-time. Machine learning was also applied to distinguish between earthquakes and microtremors (H. Liu et al., 2022) and to phase picking (Mousavi et al., 2020; W. Zhu et al., 2022). Current solutions involve convolution neural networks (Khan & Kwon, 2022; W. Zhu et al., 2022), generative adversarial networks (H. Liu et al., 2022), and transformers (Mousavi et al., 2020) applied to seismic waves. Sentinel-1 imagery has already been tested in earthquake analysis without involving automatic systems (Funning & Garcia, 2019). Specifically, the reported manual analysis highlights the feasibility of earthquake detection using Sentinel-1 data. The authors of Funning and Garcia, 2019 managed to identify approximately 70% of considered earthquakes. Compared to the manual analysis, we do not need to apply preprocessing steps for water bodies or atmospheric correction, and, more importantly, we propose completely automated ML-based approaches. Nevertheless, manual analysis in Funning and Garcia, 2019 involves correlation, which fails to understand non-linear relationships in data, while in our study, we employed many non-linear estimators. Moreover, we collected a time series of images to overcome the possible limitation of single image analysis, opening the possibility of leveraging changes instead of simply analyzing the content of single post-event images. Another difference is that we propose automatic solutions to analyze earthquakes from Sentinel-1 images instead of employing waves or manual solutions. Finally, we address several tasks, including detection and magnitude estimation.

SOURCES DESCRIPTION

In this section, we present the sources of our dataset, which includes Sentinel-1 GRD products and ground truth annotations from the International Seismological Centre (ISC).

Sentinel-1 mission

Sentinel-1 (Torres et al., 2012), with its Synthetic Aperture Radar (SAR), proved its effectiveness in a wide range of applications, and it is well suited for monitoring land changes due to its high revisit time and the capacity to see through clouds. It uses a single C-band operating at a center frequency of 5.405 GHz. SAR supports different acquisition modes in single and dual polarization: Stripmap (SM), Interferometric Wide swath (IW), Extra-Wide swath (EW), and Wave (WV). IW mode is mainly employed in land areas, WV in open oceans, EW in coastal areas, and SM for small islands. This work focused on land earthquakes, so we employed the IW mode, which provides VV and VH polarizations. Sentinel-1 products are available in four different formats: Level-0 Raw, Level-1 Ground Range Detected (GRD), Level-1 Single Look Complex (SLC), and Level-2 Ocean (OCN). We employed the Level-1 GRD products at two resolutions: 10 × 10 (high) and 40 × 40 (medium). To get all possible information, we kept only high-resolution samples.

ISC Bulletin

The International Seismological Centre (ISC) provides access to a database of all known and suspected earthquakes from more than 130 seismological agencies around the world (Willemann & Storchak, 2001). The ISC Bulletin ("International Seismological Centre (2023), On-line Bulletin", 2023) contains data from 1900 to the present (2023-11-19). The Reviewed ISC Bulletin, manually checked by ISC analysts and relocated (when there is sufficient data), is typically 24 months behind real-time and is currently up to 2021-11-01. The review procedure is applied to

all earthquakes with unknown magnitude or greater than 3.5. The procedure ensures the hypocenters are coherent with the origin region and that any regional and magnitude data are not missing. According to ISC, around 20% of the entire database is reviewed. The bulletin provides various information regarding each earthquake, such as hypocenter (Bonda´r & Storchak, 2011), the timestamp of the event, and magnitude (Di Giacomo & Storchak, 2016). Using this information, we can select an area of interest, know when to sample data related or not related to earthquakes, and try to understand different aspects of the events. The magnitude scale employed in this analysis and the ones directly provided by ISC is the mb scale, which can be used for medium entity earthquakes without encountering saturated values (Howell Jr, 1981). The following section describes how we use them to create tri-temporal time-series and ground truth annotations.

DATASET

Table 1. Dataset Summary

Earthquakes	155
Temporal window	2018-2021
Image channels	2 (VV and VH)
Area size	$20km \times 20km$
Temporal difference	1-13 days
Time series length	3
Patch size	512 × 512
Magnitudes	> 4 mb

We collected the epicenters of all earthquakes from 2018 until 2021 that were reviewed by ISC. We kept the ones defined as *known* events. We retained only the earthquakes that could create some visible effects, filtering the ones with mb magnitude greater than 4. We queried the Copernicus system of the European Space Agency to collect samples from Sentinel-1. For each epicenter longitude-latitude coordinates (E_{lg}, E_{lt}), we collected the area from $(E_{lg} - 0.1, E_{lt} + 0.1)$ to $(E_{lg} + 0.1, E_{lt} - 0.1)$. This creates an area around the epicenter of $\approx 20km \times 20km$. We have also constrained the image selection to IW mode samples. We collected two samples for each earthquake, one before (we call it "pre-event") and one after (we call it "post-event") the event. Since the revisit time of Sentinel-1 is 12 days, we created a temporal window of 13 days before and 13 days after the earthquake. We then selected the two images temporally nearest the event, excluding the day of the incident. We have excluded epicenters for which we do not have all the two samples. Furthermore, we collected an extra sample of the same area in a temporal window from 25 days before to 13 days before the earthquake, selecting the furthest day from the "pre-event" sample to avoid excessive similarities (we call it "neutral"). The availability of the third image for each earthquake helps balance the data distribution when considering the tasks based on the bitemporal time series in the following section.

Figure 1 depicts the four temporal windows considered for each earthquake. To summarize, three Sentinel-1 images are collected for each earthquake: one in the post-event, one in the pre-event, and one in the neutral temporal window.



Figure 1. Temporal windows of collected samples

Each sample was preprocessed using SAR backscatter with the sigma0 ellipsoid coefficient. In this way, we collected 155 different events. For each earthquake, we provide three images (neutral, pre-event, post-event), the magnitude expressed in mb, the epicenter coordinates, and the depth (used to identify the hypocenter), when available.

In Figure 2, it is possible to see geographical distribution around the globe. Americas, East Asia, and Oceania are best represented, while Africa has the fewest samples. The identified epicenters cover coastal areas, hinterlands, and islands.

In Figure 3a, you can see the distribution of the magnitudes. We collected sufficient samples to approximate the distribution using a normal with $\mu = 5.64$, $\sigma = 0.44$. The collected values can be considered non-saturated

according to Howell Jr, 1981, which poses the upper limit of the scale to $mb_{sat} \approx 8.1$.

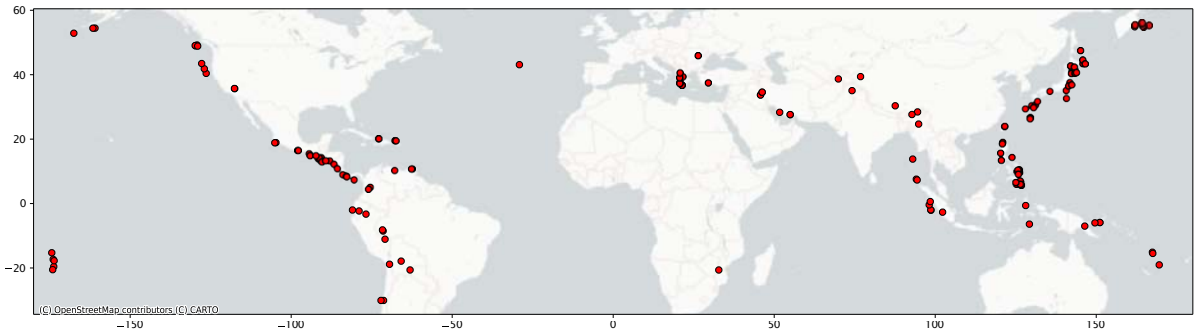
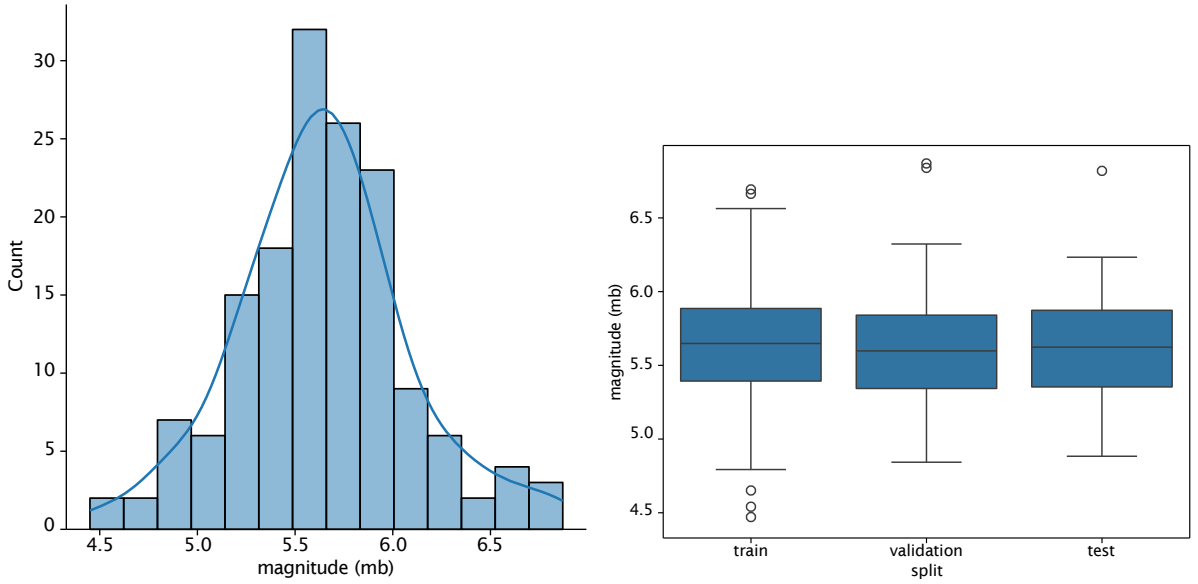


Figure 2. Earthquakes epicenters around the globe

We patched the samples from the center to obtain images of size 512 × 512 × 2, which are more tractable by machine learning models. In this way, we created 1906 patches. We reported a dataset summary in Table 1, highlighting the main features.



(a) Magnitudes distribution

(b) Magnitude distribution over splits

Figure 3. Magnitude distributions

Train, Validation and Test Splits

To create representative splits, we divided the dataset into 70%-15%-15% sets for train, validation, and testing, respectively. We decided to create the splits according to the distribution of the magnitudes. All the images (neutral, pre-event, and post-event) associated with the same earthquake event are in the same set/split. We sample from the original dataset with a uniform distribution. We evaluated the magnitude distribution using the two-sample Kolmogorov-Smirnov test to ensure similar distributions between train-validation and train-test. We obtain a p-value of 0.985 between train-validation and 0.995 between train-test sets. The obtained distributions in Figure 3b highlight the similarities between the three sets in terms of quantiles.

Tasks

Several tasks related to earthquake monitoring can be solved using the data we collected, and we are sharing. Among the others, we suggest using our data to handle the following emergency earthquake-related tasks:

1. Earthquake detection with bitemporal image time series
2. Magnitude regression with bitemporal image time series

3. Hypocenter or epicenter regression with bitemporal image time series
4. Change detection with bitemporal image time series
5. Earthquake detection with a single image
6. Magnitude regression with a single image
7. Hypocenter or epicenter regression with a single image

In the following, we formalize each type of task regardless of the input image time series size (the input can be either a single image or a time series composed of two consecutive images).

Earthquake Detection

The earthquake detection tasks (Tasks 1 and 5) can be formulated as follows. We have a set of time-series T, where each time-series $T_s \in \mathrm{T}$ is composed of N images, with $N \leq 2$, of size $W \times H \times C$ related to the same spatial area S at timesteps $\{T_1, \ldots, T_N\}$ and a ground truth value $G_t \in \{0, 1\}$ (where 1 indicate S was affected by an earthquake, 0 otherwise). Given a training dataset D_{tr}, composed of a set of pairs (T_s, G_t), we train a machine learning model
M. Given a test dataset D_{ts}, composed of a set of T_s, we can predict G_t for each sample T_s using M. This can be framed as a supervised classification task.

Magnitude Regression

The magnitude regression tasks (Tasks 2 and 6) can be formulated as follows. We have a set of time-series T, where each time-series $T_s \in \mathrm{T}$ is composed of N images, with $N \leq 2$, of size $W \times H \times C$ related to the same spatial area S at different timesteps $\{T_1, \ldots, T_N\}$ and a ground truth value $G_t \in \{0 \ldots M_m\}$ (where M_m is the maximum value for the given magnitude scale). Given a training dataset D_{tr}, composed of a set of pairs (T_s, G_t), we train a machine learning model M. Given a test dataset D_{ts}, composed of a set of T_s, we can regress G_t for each sample T_s using
M. This regression task is a supervised one.

Epicenter or Hypocenter Regression

The Epicenter or Hypocenter regression tasks (Tasks 3 and 7) can be formulated as follows. We have a set of time-series T, where each time-series $T_s \in \mathrm{T}$ is composed of N images, with $N \leq 2$, of size $W \times H \times C$ related to the same spatial area S at different timesteps $\{T_1, \ldots, T_N\}$ and a ground truth composed by the coordinates (x, y, d) (where x and y are the epicenter coordinates and d is the depth of the hypocenter) and a binary label $G_t \in \{0, 1\}$ (1 if the area S contains the hypocenter of the earthquake, otherwise 0). Given a training dataset D_{tr}, composed of a set of tuples (T_s, G_t, x, y, d), we train a machine learning model M. Given a test dataset D_{ts}, composed of a set of T_s, we can predict (G_t, x, y, d) for each sample T_s using M. If S is affected by an earthquake but does not contain the hypocenter, it should recognize that the tuple (x, y, d) cannot be determined. The task is supervised and combines classification (presence or absence of the epicenter) and regression (coordinates of the hypocenter).

Change Detection

Change detection (Task 4) can be formulated as follows. We have a set of time-series T, where each time-series $T_s \in \mathrm{T}$ is composed of 2 images of size $W \times H \times C$ related to the same spatial area S at different timesteps $\{T_1, T_2\}$. We want to generate a binary mask M_b of size $W \times H$ for each T_s. The value 1 indicates the pixel is affected by a substantial change; otherwise, it is set to 0. The task is an unsupervised semantic segmentation task. Given a training dataset D_{tr}, composed of a set of T_s, we train a machine learning model M. Given a test dataset D_{ts}, composed of a set of T_s, we can generate M_b for each sample T_s using M.

The previous formulations are general, and the mentioned tasks can be solved using single images (degenerate time series) or bi-temporal time series composed of two images.

EXPERIMENTS

In this section, we present the results obtained for a subset of the proposed tasks, which are Tasks 1, 2, 5 and 6. We have solved the most promising tasks to evaluate the quality of the collected data and the feasibility of the more relevant proposed tasks.

We employed four deep-learning models designed for low-resource devices: MobileNetV2 (Sandler et al., 2018) and ConvNextV2-Atto (Woo et al., 2023), which are convolutional neural networks (CNN), and MiT-B0 (Xie

et al., 2021) and MobileViTV2-1.0 (Mehta & Rastegari, 2022), which are vision transformers (ViT). We also considered some classical shallow learning models: Support Vector Machines (SVM) with 3rd-grade polynomial and Radial Basis Function kernels and Random Forest (RF). We have chosen these solutions because SVMs can exploit non-linear relations, while RFs, being an ensemble, are more robust.

Experimental Settings

This section reports all settings for the employed models and the input datasets.

Input data

Single Image. In this setting, the input data is a collection of single images (i.e., a collection of degenerate time series, each one composed of a single image). We use all the post-event images (as representative of areas already affected by an earthquake) and all the pre-event ones (as representative of areas unaffected by an earthquake). This corresponds to a balanced distribution of the two classes under analysis (affected/unaffected by an earthquake). In this setting, we do not consider the neutral images (see Figure 1). Each image has shape 512 × 512 × 2.

Bi-temporal Time Series. This second setting considers as input a collection of bitemporal time series of images. When dealing with bitemporal time series of images, we use samples of shape 512 × 512 × 4 concatenating the two components of the time series of shape 512 × 512 × 2 along the channel axis (this approach allows using machine learning models not designed for time series). Each time series is related to a specific earthquake and comprises a pre-event and a post-event image (example of affected area class object) or a neutral and a pre-event sample (example of unaffected area class object). This ensures a balanced distribution of positive and negative examples for the two classes.

Models

Deep Learning Models. We trained convolutional neural networks for six epochs while vision transformers for ten epochs. This was done due to the slower learning of transformers. The batch size is 16. We used an AdamW optimizer with a cosine-annealing learning rate scheduler with a warmup of 0.1 of the total training steps. The learning rate starts from 0.0001 for CNNs and from 0.001 for ViTs. All models are initialized without any pre-trained weights. The loss is Mean Squared Error (MSE) for regression tasks and Cross-Entropy loss for classification.

Shallow Models. We trained the Support Vector Machines with 3rd-degree polynomial and RBF kernels. Due to the high number of features, we applied PCA before using classical models. In this way, the size of the features of each sample is reduced to ≈ 2000.

Metrics

We analyzed the results in terms of accuracy for classification tasks since the class distribution is balanced (≈ 50% of positive and negative samples). Accuracy is defined as:

$$Accuracy = \frac{CC}{N} \tag{1}$$

where CC is the number of samples predicted correctly, and N is the number of samples we asked for a prediction.

We analyzed the results regarding Mean Absolute Error (MAE) for the regression tasks. Still, we also evaluated the accuracy in identifying the occurrence of an earthquake thresholding the predicted magnitude (if magnitude < 1, then no earthquake occurred; otherwise, an event affected the area). Mean Absolute Error is defined as:

$$\text{MAE} = \frac{1}{N} \sum_{i}^{N} |y_i - \hat{y}_i| \tag{2}$$

where N is the number of samples, y_i is the ground truth magnitude, and $\hat{y}_i$ is the predicted magnitude for the ith sample.

We are also interested in the resource consumption of the presented models, so we report the number of parameters for deep learning models, inference time in seconds, and Mega Floating Points Operation Per seconds (MFLOPs) computed on an Intel(R) Core(TM) i9-10980XE CPU.

Earthquake Detection with Bitemporal Time Series

We report in Table 2 the results for Task 1. The Random Forest Classifier is the best shallow model in terms of accuracy, providing reasonably high accuracy with a few FLOPs.

Regarding deep learning models, MobileViTV2 struggles to compete with the others, obtaining −20% accuracy and 2× FLOPs compared to Mit-B0. The best accuracy results are obtained by the CNNs, which are also less resource-hungry than transformers. MobileNet achieves the best result (94.72%) with fewer FLOPs than the other deep learning models. The differences between this model and the RFC classifier regarding FLOPs and accuracy are evident. Hence, deep learning models are needed to achieve high accuracy values at the cost of a more elevated resource consumption (two orders of magnitude more FLOPs than RFC). However, we notice that the "smallest" model (fewer parameters) of each deep learning category performs better than the "bigger" model of the same category. Specifically, MobileNetV2 (CNN) is more accurate than ConvNextV2 (CNN) and MiT-B0 (ViT) is more accurate than MobileVitV2 (ViT). Hence, smaller and less resource-demanding deep learning models are better for solving this task.

Table 2. Performance of models for earthquake detection with bitemporal time-series

Model	Params	Accuracy ↑	Time (s) ↓	MFLOPs ↓
SVC (RBF kernel)	-	0.6341	0.3640	1.0486
SVC (Poly kernel)	-	0.5440	0.3138	1.0486
RFC	-	0.7241	0.3130	1.0508
MobileNetV2 (CNN)	2.2M	0.9472	0.1109	207.5949
MiT-B0 (ViT)	3.4M	0.8865	0.0909	430.7062
ConvNextV2 (CNN)	3.7M	0.9374	0.1003	373.7727
MobileViTV2 (ViT)	4.9M	0.6536	0.1746	964.5956

Magnitude Regression with Bitemporal Time Series

In Table 3, we report the results obtained for Task 2. Classical solutions like Support Vector Regressors and Random Forest Regressors struggle to compete with deep learning models. In this case, the SVR with RBF kernel gets the best results in terms of accuracy among shallow machine learning models, despite RFR being the best in MAE. This means SVR can better distinguish between positive and negative samples, while RFR tends to approximate the magnitude better. Another critical factor is the accuracies of the shallow models, which are all near 0.5, being generally incapable of making distinctions between the classes.

Deep learning models are also more accurate (in terms of accuracy and MAE) for this task. Deep learning models provide higher accuracies and MAEs, distinguishing affected and unaffected areas. MiT-B0, which achieves worse MAE than CNNs (ConvNextV2 and MobileNetV2), gets an accuracy comparable to ConvNextV2 and MobileNetV2. Hence, Mit-B0, ConvNext, and MobileNet have the same ability to distinguish between earthquake-affected/unaffected areas. However, ConvNextV2 and MobileNetV2 estimate the magnitude values better. The performances of MobileViTV2 are not as satisfactory as the other deep-learning models. Similarly to the previous task, MobileNetV2 obtains the best performance under all metrics among the deep learning models despite being the "smallest" deep learning model. Furthermore, MobileNetV2 (CNN) is again more accurate (accuracy) and precise (MAE) than ConvNextV2 (CNN), and MiT-B0 (ViT) is more accurate and precise than MobileVitV2 (ViT). Hence, smaller and less resource-demanding deep learning models are also better for solving this task.

Table 3. Performance of models for magnitude regression with bitemporal time-series

Model	Params	MAE ↓	Accuracy ↑	Time (s) ↓	MFLOPs ↓
SVR (RBF kernel)	-	2.2368	0.5519	0.3640	1.0486
SVR (Poly kernel)	-	2.6015	0.5440	0.3138	1.0486
RFR	-	1.9930	0.5440	0.3130	1.0508
MobileNetV2 (CNN)	2.2M	0.5456	0.9374	0.1109	207.5949
MiT-B0 (ViT)	3.4M	0.8496	0.9276	0.0909	430.7062
ConvNextV2 (CNN)	3.7M	0.6494	0.9315	0.1003	373.7727
MobileViTV2 (ViT)	4.9M	1.7612	0.7378	0.1746	964.5956

Earthquake Detection with a Single Image

Table 4 reports the results for Task 5. It shows the models struggle to discriminate between areas affected by earthquakes and the non-affected ones. The benefits in reducing the computation costs compared to the models trained on bitemporal input data are evident only for shallow models (≈ 0.5× FLOPs than bitemporal settings), while the deep learning models gain is not worth the decrease in terms of accuracy (compare Table 4 with Table 2).

The best shallow model to solve Task 5 is SVC with RBF kernel, which achieves 0.60 of accuracy in a few FLOPs. Transformers get worse results than this model (−6% in the worst case). MobileNetV2 provides the best results (with +10% than SVC with RBF kernel), followed by ConvNextV2.

MobileNetV2 with single images achieves −24% in terms of accuracy than bitemporal settings (95%), and the results are comparable to those of Random Forest (RFC) with bitemporal input data (72%).

A consistent performance decrease using only one image was expected, as shown in the manual analysis of interferograms of Funning and Garcia, 2019 in which the authors highlighted that earthquake detection is feasible by looking at the differences between the two phases.

Table 4. Performance of models for earthquake detection with a single image

Model	Params	Accuracy ↑	Time (s) ↓	MFLOPs ↓
SVC (RBF kernel)	-	0.6047	0.1488	0.524288
SVC (Poly kernel)	-	0.5440	0.1472	0.524288
RFC	-	0.5186	0.1655	0.526554
MobileNetV2 (CNN)	2.2M	0.7104	0.0862	203.3172
MiT-B0 (ViT)	3.4M	0.5440	0.0710	424.8801
ConvNextV2 (CNN)	3.7M	0.6419	0.0609	370.8897
MobileViTV2 (ViT)	4.9M	0.5988	0.1171	960.3690

Magnitude Regression with a Single Image

The results in Table 5 for Task 6 show, as expected from the previous experiments, that the models are not well capable of identifying affected areas using one single image. We can conclude from the reported accuracies that this is common for both shallow and deep learning models.

The differences between the used models in terms of MAE and accuracy are less evident than before. Shallow models align with previous performances for Task 2. Conversely, deep learning models decrease their mean performance of −30% in terms of accuracy and +1.3 in MAE compared to the bitemporal setting.

MobileNetV2 gets the best results in terms of MAE but does not perform well in distinguishing between affected and unaffected areas. The decrease in accuracy compared to Task 5 is −16%. ConvNextV2 has the best accuracy, while MiT-B0 is the most balanced solution in both metrics. MobileViTV2 struggles to compete with shallow models in both metrics, too.

Table 5. Performance of models for magnitude regression with a single image

Model	Params	MAE ↓	Accuracy ↑	Time ↓	MFLOPs ↓
SVR (RBF kernel)	-	2.4088	0.5519	0.1488	0.524288
SVR (Poly kernel)	-	2.5960	0.5440	0.1472	0.524288
RFR	-	2.6204	0.5440	0.1655	0.526554
MobileNetV2 (CNN)	2.2M	2.1595	0.5734	0.0862	203.3172
MiT-B0 (ViT)	3.4M	2.1821	0.5832	0.0710	424.8801
ConvNextV2 (CNN)	3.7M	2.2353	0.5851	0.0609	370.8897
MobileViTV2 (ViT)	4.9M	2.4645	0.5440	0.1171	960.3690

Preliminary Integration of a Global Seismic Hazard Map

The integration of the right in-domain information could be beneficial in addressing the proposed tasks. In our preliminary assessment, we evaluated the effectiveness of integrating a global seismic hazard map (Johnson et al.,

2023) as input. Specifically, we performed preliminary experiments by adding an extra channel associated with the seismic hazard information to the bi-temporal images. According to the preliminary results, this additional information does not provide any benefits in earthquake detection and magnitude regression. In the future, we will investigate other in-domain sources as well as other ways of integrating domain-specific information in the training of the models.

LIMITATIONS

The study employed only visual features without any external knowledge. Including other historical features and employing domain knowledge (e.g., geophysics, optics, and earth sciences) could strengthen the results on more complex tasks like hypocenter regression, which could benefit from extra information like geological plaques.

We employed single images and bi-temporal time series, although a longer time series could be investigated to include more historical information about the analyzed areas.

CONCLUSION AND FUTURE DIRECTIONS

Our experiments show the possibility of exploiting machine learning techniques to understand earthquakes using the visual components collected by the SAR instrument with satisfactory results.

We explored the usage of Sentinel-1 imagery, which is easily available to the public and for which there were studies proving effectiveness, although a different source (e.g., from drones and airplanes) or multi-source approach could be employed to adopt less expensive instrumentation.

We employed low-resource solutions (which are generally "small" models in terms of parameters) to show the unnecessary need for complex and bigger models. Results show that smaller deep-learning models are more accurate and precise than the bigger ones, contrasting expectations. Designing ad-hoc low-resource networks is an underexplored direction that should be addressed in remote sensing and hazard management.

In future works, we plan to investigate the remaining proposed tasks, the design of new ad-hoc low-resource models, and the inclusion of in-domain knowledge.

REFERENCES

Alam, F., Imran, M., & Ofli, F. (2017). Image4act: Online social media image processing for disaster response. *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, 601–604.

Asokan, A., & Anitha, J. (2019). Change detection techniques for remote sensing applications: A survey. *Earth Science Informatics*, *12*, 143–160.

Bonafilia, D., Tellman, B., Anderson, T., & Issenberg, E. (2020). Sen1floods11: A georeferenced dataset to train and test deep learning flood algorithms for sentinel-1. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

Bonda´r, I., & Storchak, D. (2011). Improved location procedures at the international seismological centre. *Geophysical Journal International*, *186*(3), 1220–1244.

Cambrin, D. R., Colomba, L., & Garza, P. (2023). Cabuar: California burned areas dataset for delineation [software and data sets]. *IEEE Geoscience and Remote Sensing Magazine*, *11*(3), 106–113. https://doi.org/10.1109/MGRS.2023.3292467

Chandrakumar, C., Prasanna, R., Stephens, M., Tan, M. L., Holden, C., Punchihewa, A., Becker, J. S., Jeong, S., & Ravishan, D. (2023). Algorithms for detecting p-waves and earthquake magnitude estimation: Initial literature review findings.

Daly, S., & Thom, J. A. (2016). Mining and classifying image posts on social media to analyse fires. *ISCRAM*, 1–14.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. *2009 IEEE conference on computer vision and pattern recognition*, 248–255.

Di Giacomo, D., & Storchak, D. A. (2016). A scheme to set preferred magnitudes in the isc bulletin. *Journal of Seismology*, *20*, 555–567.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., et al. (2012). Sentinel-2: Esa's optical high-resolution mission for gmes operational services. *Remote sensing of Environment*, *120*, 25–36.

Funning, G. J., & Garcia, A. (2019). A systematic study of earthquake detectability using sentinel-1 interferometric wide-swath data. *Geophysical Journal International*, *216*(1), 332–349.

Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(7), 2217–2226.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

Howell Jr, B. (1981). On the saturation of earthquake magnitudes. *Bulletin of the Seismological Society of America*, *71*(5), 1401–1422.

International seismological centre (2023), on-line bulletin. (2023). https://doi.org/https://doi.org/10.31905/D808B830

Ji, S., Wei, S., & Lu, M. (2018). Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. *IEEE Transactions on geoscience and remote sensing*, *57*(1), 574–586.

Johnson, K., Villani, M., Bayliss, K., Brooks, C., Chandrasekhar, S., Chartier, T., Chen, Y.-S., Garcia-Pelaez, J., Gee, R., Styron, R., Rood, A., Simionato, M., & Pagani, M. (2023). Global seismic hazard map. https://doi.org/10.5281/ZENODO.8409647

Justice, C., Townshend, J., Vermote, E., Masuoka, E., Wolfe, R., Saleous, N., Roy, D., & Morisette, J. (2002). An overview of modis land data processing and product status. *Remote sensing of Environment*, *83*(1-2), 3–15.

Khan, I., & Kwon, Y.-W. (2022). P-detector: Real-time p-wave detection in a seismic waveform recorded on a low-cost mems accelerometer using deep learning. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5.

Liu, H., Li, S., & Song, J. (2022). Discrimination between earthquake p waves and microtremors via a generative adversarial network. *Bulletin of the Seismological Society of America*, *112*(2), 669–679.

Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A convnet for the 2020s. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11976–11986.

Long, Z., McCreadie, R., & Imran, M. (2023). Crisisvit: A robust vision transformer for crisis image classification. https://doi.org/http://dx.doi.org/10.59297/SDSM9194

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Mu¨ller-Wilm, U., & Gascon, F. (2017). Sen2cor for sentinel-2. *Image and Signal Processing for Remote Sensing XXIII*, *10427*, 37–48.

McCreadie, R., Buntain, C., & Soboroff, I. (2019). Trec incident streams: Finding actionable information on social media.

McCreadie, R., Buntain, C., Soboroff, I., & Ellis, A. (2023). Crisisfacts: Buidling and evaluating crisis timelines. *Proceedings of the 20th International Conference on Information Systems for Crisis Response and Management*. https://doi.org/http://dx.doi.org/10.59297/JVQZ9405

Mehta, S., & Rastegari, M. (2021). Mobilevit: Light-weight, general-purpose, and mobile-friendly vision transformer. *arXiv preprint arXiv:2110.02178*.

Mehta, S., & Rastegari, M. (2022). Separable self-attention for mobile vision transformers. *arXiv preprint arXiv:2206.02680*.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications*, *11*(1), 3952.

Nguyen, D. T., Ofli, F., Imran, M., & Mitra, P. (2017). Damage assessment from social media imagery data during disasters. *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, 569–576.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4510–4520.

Shao, Z., Wu, W., Wang, Z., Du, W., & Li, C. (2018). Seaships: A large-scale precisely annotated dataset for ship detection. *IEEE transactions on multimedia*, *20*(10), 2593–2604.

Sumbul, G., Charfuelan, M., Demir, B., & Markl, V. (2019). Bigearthnet: A large-scale benchmark archive for remote sensing image understanding. *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, 5901–5904.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., et al. (2012). Gmes sentinel-1 mission. *Remote sensing of environment*, *120*, 9–24.

Wang, W., Xie, E., Li, X., Fan, D.-P., Song, K., Liang, D., Lu, T., Luo, P., & Shao, L. (2021). Pyramid vision transformer: A versatile backbone for dense prediction without convolutions. *Proceedings of the IEEE/CVF international conference on computer vision*, 568–578.

Wang, Y., Braham, N. A. A., Xiong, Z., Liu, C., Albrecht, C. M., & Zhu, X. X. (2022). Ssl4eo-s12: A large-scale multi-modal, multi-temporal dataset for self-supervised learning in earth observation. *arXiv preprint arXiv:2211.07044*.

Weber, E., Papadopoulos, D. P., Lapedriza, A., Ofli, F., Imran, M., & Torralba, A. (2022). Incidents1m: A large-scale dataset of images with natural disasters, damage, and incidents. *IEEE transactions on pattern analysis and machine intelligence*, *45*(4), 4768–4781.

Willemann, R. J., & Storchak, D. A. (2001). Data collection at the international seismological centre. *Seismological Research Letters*, *72*(4), 440–453.

Woo, S., Debnath, S., Hu, R., Chen, X., Liu, Z., Kweon, I. S., & Xie, S. (2023). Convnext v2: Co-designing and scaling convnets with masked autoencoders. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16133–16142.

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., Allen, R. G., Anderson, M. C., Belward, A. S., Cohen, W. B., et al. (2019). Current status of landsat program, science, and applications. *Remote sensing of environment*, *225*, 127–147.

Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). Segformer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*, *34*, 12077–12090.

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., et al. (2020). Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment*, *241*, 111716.

Zhu, W., Tai, K. S., Mousavi, S. M., Bailis, P., & Beroza, G. C. (2022). An end-to-end earthquake detection method for joint phase picking and association using deep learning. *Journal of Geophysical Research: Solid Earth*, *127*(3), e2021JB023283.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307



ISCRAM 2024
MÜNSTER, GERMANY

Volunteers in Crisis and Emergency Management

Track Chairs

Niki Matinrad, Linköping University

Sofie Pilemalm, Linköping University University of Agder

Caroline Rizza, I3-Telecom Paris, Institut Polytechnique de Paris

Ophélie Morand, I3-Telecom Paris, Institut Polytechnique de Paris

Kayvan Yousefi Mojir, University of Borås

Managing Volunteers in Disaster Situations - An Overview of Models for Decision-Makers

Florentina Hager*
University of Twente
f.hager@utwente.nl

Melanie Reuter-Oppermann
University of Twente
m.n.reuter-oppermann@utwente.nl

ABSTRACT

Volunteers are a valuable resource in disaster management. Yet to be useful, decision-makers need to manage this resource properly. However, managing volunteers differs from managing traditional work forces in multiple aspects: volunteers can spontaneously decide to offer their help, may have different skills and preferences for different tasks and can leave the system at any time. Several models for optimising volunteer management in disaster scenarios have been presented in the past. To provide an overview of existing approaches, this article provides a structured literature review, discusses the challenges of integrating these models into decision support systems and identifies future research directions. In total, 31 articles were identified, of which more than half were published in the last five years. Most authors addressed the issue of assigning volunteers to tasks, while little attention was paid to predicting the number of volunteers and optimising the composition of volunteer teams.

Keywords

Volunteers, decision support, disaster management, operational research, literature review

INTRODUCTION

When a disaster strikes, many different resources, such as medical or technical equipment and human resources, need to be available as quickly as possible. Thereby volunteers can play a significant role. For instance, according to a study conducted by the Fritz Institute in Indonesia after the 2004 tsunami, 90% of the survivors surveyed reported being rescued by volunteers (Fritz Institute 2005). If volunteers had their own sovereign nation, it would rank as the second most populous country in the world with its economy being one of the ten largest in the world (Salamon et al. 2011).

There are several definitions of volunteers. Most of them have in common that volunteers do not work for a salary but for free or only receive remuneration in the form of food and are not obliged to volunteer (Sampson 2006). Among volunteers, a second distinction is made in terms of affiliation. While some volunteers belong to an organisation such as the Red Cross, where they may receive training or similar, people may also decide to volunteer spontaneously, especially after a disaster. Therefore, these volunteers are sometimes referred to as "spontaneous" volunteers. However, this term implies that volunteers with affiliation are not spontaneous. Hence, we agree with Daddoust et al. (2021) that they should rather be called "unaffiliated emergent" volunteers. However, to be consistent with the literature, we will refer to these volunteers as spontaneous volunteers (SV). For a more detailed discussion on the definition of volunteers and SV, we refer the reader to Sampson (2006) and Paciarotti and Cesaroni (2020).

SV can be a useful resource in the aftermath of a disaster. They may be quickly available (Daddoust et al. 2021) and provide additional work force and possibly also additional skills (L. Fernandez et al. 2006) for little to no cost (Daddoust et al. 2021). Even without specialised skills, they can take on simple tasks such as making sandbags in the event of flooding or distributing food to relieve the burden on emergency management staff allowing them to focus on more specialised activities (Sauer et al. 2014).

*corresponding author

However, with a lack of an underlying organisation that exists for affiliated volunteers, the deployment of SV can also pose "serious issues and risks" (L. S. Fernandez et al. 2006). In the worst case, they can even cause additional problems after a disaster (L. Fernandez et al. 2006, S.-Q. Chen and Bai 2023) by obstructing or interfering with emergency response personnel and/or putting themselves or the victims in danger by acting according to what they think is necessary (Whittaker et al. 2015). Despite all the risks associated with the involvement of SV, it is also not an option not to involve them in disaster relief when they are available. Forgoing this additional labour force means a loss of resources, which can also distort the public perception of disaster response (L. Fernandez et al. 2006). In addition, the rejection of SV can also reduce future willingness to volunteer (Falasca, Zobel, and Fetter 2009, Daddoust et al. 2021). Hence, emergency managers should be prepared to manage volunteers in order to utilise their potential as efficiently as possible while avoiding additional risks and threads. According to research, organisations should have a centralised system to administer SV and assign tasks to them individually or in groups (L. Fernandez et al. 2006). Such systems should be able to respond to changing situations during disasters and be applied dynamically "in an on-going process", as well as anticipate needs (L. Fernandez et al. 2006).

Models developed in the field of operations research can provide decision-makers with a solution for the management of volunteers and the most efficient use of their workforce. For practical application, integrating these models into decision support systems (DSS) is important to facilitate the decision-making process. This review aims to provide an overview of existing models for the management of SV in disaster situations. The objective of this work is twofold: on the one hand, we aim to point out gaps and future directions for volunteer management in disasters. On the other hand, we want to facilitate the design of future DSS by providing an overview of the models and the data requirements that such a system should fulfil. In this work, we focus on physical volunteers who do not (only) collect data, but "offer their workforce and support at the disaster response sides within their range" (Betke 2018). In addition, we only consider volunteers in disaster situations and exclude regular emergency situations.

The remainder of this paper is structured as follows: In the next section, we provide an overview of the related literature. We then describe the methodology for the structured literature review and present the results, including an outlook on future research topics, before concluding the paper with a summary.

RELATED LITERATURE

The problem of labour scheduling has been extensively studied in the past (Van Den Bergh et al. 2013). Although volunteers can be understood as a type of labour, some assumptions made in traditional labour allocation (TLA) do not necessarily apply to voluntary labour allocation (VLA) (Sampson 2006). Following the most common definitions presented in the previous section, volunteers generally do not receive compensation, except perhaps per diems. Therefore, the overall goal shifts from minimising labour costs to maximising the tasks completed (Krstikj et al. 2021). Another key difference between TLA and VLA is the reliability of the workforce. Without significant remuneration or an employment contract, volunteers have no obligation to be available for work (Whittaker et al. 2015). Especially when dissatisfied with the assigned tasks or the amount of idle time, volunteers may decide to withdraw their help (Garcia et al. 2018). These particular characteristics inherent in the management of SV have led to extensive research into the psychological, motivational, and demographic aspects of volunteering (Einolf 2018). However, despite this general interest, the management of SV has received very little attention in the field of operational research (Sampson 2006).

The additional capacity of human resources provided by volunteers can be particularly valuable during disasters. Emergency response in disaster situations differs in various aspects from everyday emergency response (Altay and Green 2006) and poses its own unique challenges. Several articles have already been published analysing the existing literature in the field of Operational Research (OR)/Management Science (MS) addressing these challenges. One of the first reviews was written by Altay and Green (2006), who reviewed the literature up to 2004. This review was followed by several other reviews, including Galindo and Batta (2013), Özdamar and Ertem (2015), Kovacs and Moshtari (2019) or more recently Farahani et al. (2020). Despite their broad focus on various aspects of disaster management, none of the reviews explicitly addresses volunteer management. This gap in the literature is partially addressed by Ozdemir et al. (2022), who conduct a structured literature review focusing on volunteer location, allocation, and assignment. However, as the articles reviewed are limited to one of the three tasks mentioned, they only cover a subset of the decisions faced by emergency managers.

In addition to the literature in the field of OR/MS, the issue of volunteer management has also received attention in the field of information systems, particularly in association with the design of DSS. To guide the development of a DSS, L. S. Fernandez et al. (2006) recommend a system-based approach. For an in-depth evaluation of Volunteer Management Systems (VMS), Schonbock et al. (2016) develop a reference model and conduct a detailed comparison of seven distinct VMS. In a broader context, Betke et al. (2023) introduce a taxonomy that provides a comprehensive overview of these systems, facilitating their classification and selection based on relevant characteristics.

Apart from articles generally supporting the design of DSS, several authors have also contributed to the development of such systems. Among others, Fajardo and Oppus (2010), Rauchecker and Schryen (2018), Razali et al. (2019) and Patrisina et al. (2021) outline approaches to building DSS in the context of disaster management. A comprehensive documentation of the entire design approach is provided by Lindner, Betke, et al. (2017), Lindner, Kuehnel, et al. (2018) and Betke (2018). With a focus on agent-based simulation, Lindner, Betke, et al. (2017) identify 25 relevant attributes of spontaneous unaffiliated volunteers. Building on this work, Lindner, Kuehnel, et al. (2018) present a conceptual model for simulating volunteer behaviour, demonstrated in an artefact in Betke (2018).

Despite the strong interconnection between these two fields, there is a limited body of literature addressing the integration of OR/MS models into DSS. Conversely, the information system literature tends to concentrate on the technical implementation of DSS, neglecting the exploration of models guiding specific decisions.

Our distinctive contribution lies in a systematic literature review providing a comprehensive overview of existing models designed to enhance decision-making regarding volunteers in disaster situations, with a particular focus on their integration into DSS. Furthermore, our review extends the scope of Ozdemir et al. (2022) beyond location, allocation, and task assignment to encompass all potential aspects decision-makers might encounter in volunteer management pre- and post-disaster. We also include models designed for slow-onset disasters, and consider the actions taken during these events as part of the post-disaster phase.

METHODOLOGY

To provide an accurate overview of the currently existing literature, we conducted a systematic literature review and present the results according to PRISMA 2020 (Page et al. 2021). We searched two databases, Scopus, and Web of Science, using the following query to search by title, abstract and keywords: *volunteer* AND (disaster* OR crisis*) AND (optimi* OR alloc* OR quanti*)* and restricted the language to English. This query led to 365 results in Scopus, and 249 results in Web of Science. Additionally, to obtain a complete list of articles related to volunteer management within the field of OR/MS, we performed a second search in Web of Science with a more general query: *volunteer* AND (disaster* OR crisis*)* and restricted the results to those articles linked to the category of "Operations Research and Management", again considering only publications written in English. This search yielded to 43 hits. We then removed 227 duplicates as well as notes and editorials, conference reviews, books, and book chapters. Eventually, 397 articles remained for screening. After reviewing titles and abstracts, we excluded papers that did not focus on volunteer management in disasters, leaving 80 publications for further screening. Out of these, five publications could not be retrieved and were therefore excluded. During the full-text screening, we included all articles that addressed decision-making related to volunteers in disaster situations, while excluding those related to non-disaster situations, such as regular emergency scenarios, or that lacked a model supporting volunteer-related decisions. Furthermore, we opted to exclude the paper authored by Gutjahr and Viswanath (2020) due to its lack of consideration of the characteristics of volunteers highlighted in the previous section. Although their assumption might be valid in certain scenarios, decision-makers adhering to such assumptions could alternatively use models developed for general work assignment problems. Based on the 29 papers initially selected, we conducted a forward and backward search to ensure the inclusion of all relevant papers. In this process, we identified two additional articles, for which we also performed a forward and backward search, where, however, no other relevant articles where identified. A visual representation of the workflow is outlined in Figure 1.

Similar to other fields of disaster management, such as casualty management (Farahani et al. 2020), the potential of volunteer management has only recently gained attention. Despite the limited number of publications, interest has remained constant with a few publications per year. Figure 2 depicts the timely distribution of the articles.

The diversity of fields related to volunteer management is also reflected in the number of different journals and proceedings in which the articles were published. In total, the 31 articles were published in 24 different journals and proceedings, ranging from OR-related journals such as the European Journal of Operational Research to journals and conferences more closely related to computer science such as the International Journal of Advanced Computer Science and Applications.

Based on the questions addressed in the articles, we have identified seven different decision areas for emergency management. These fields include predicting the number of SV needed, hiring/admitting volunteers, training volunteers to acquire new skills, pairing volunteers, locating them in different geographical areas, and allocating them to tasks and queues. Figure 3 shows the distribution[1]. The most studied area is the assignment of volunteers to tasks, with more than half of the selected articles addressing this aspect. Some papers propose an approach to the hiring/admission of volunteers, their location and allocation to queues. The aspects of prediction, training, and pairing, however, are hardly addressed in the literature. Table 1 provides a brief overview of the selected articles.

[1]Note that some articles refer to two areas and are therefore counted twice.

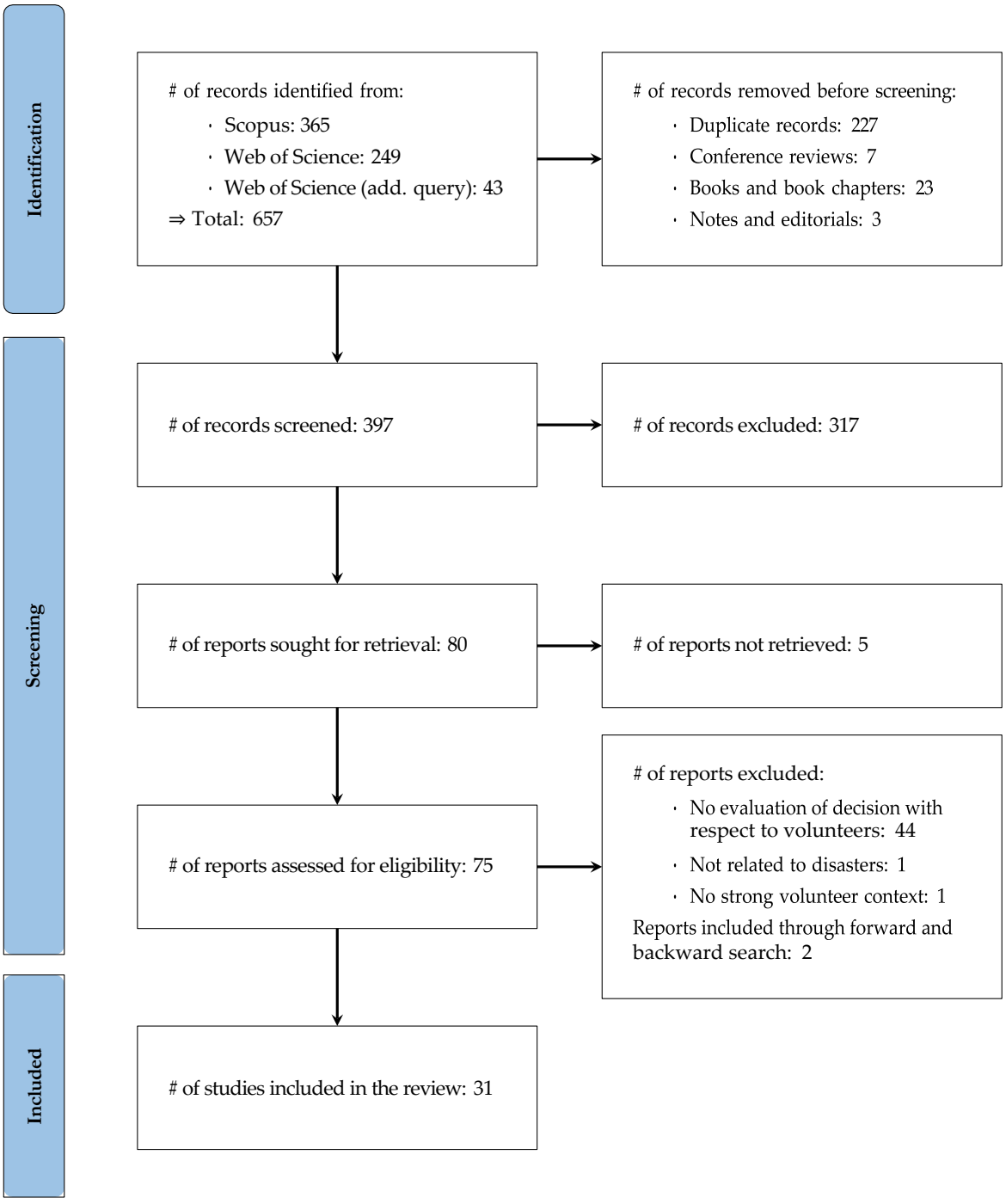


Figure 1. Review process mapped to PRISMA 2020 (following Page et al. 2021)

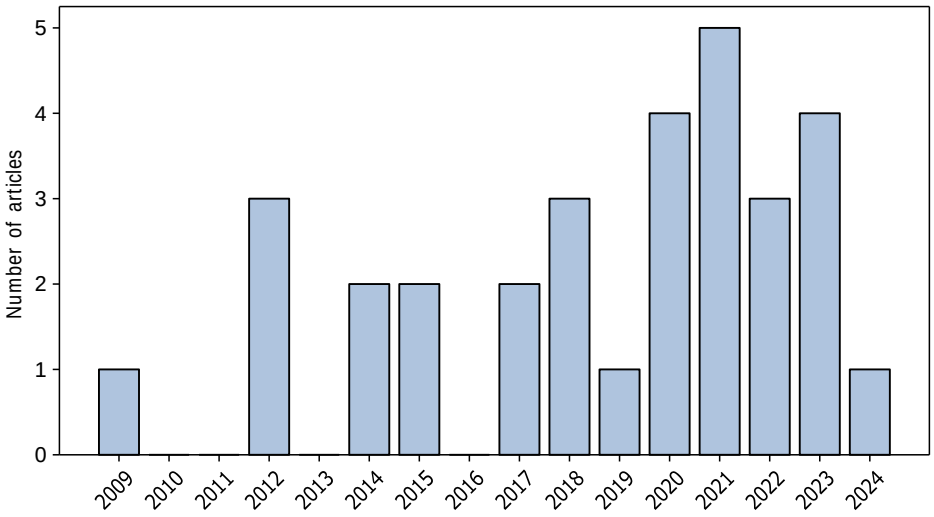


Figure 2. Timely distribution of papers

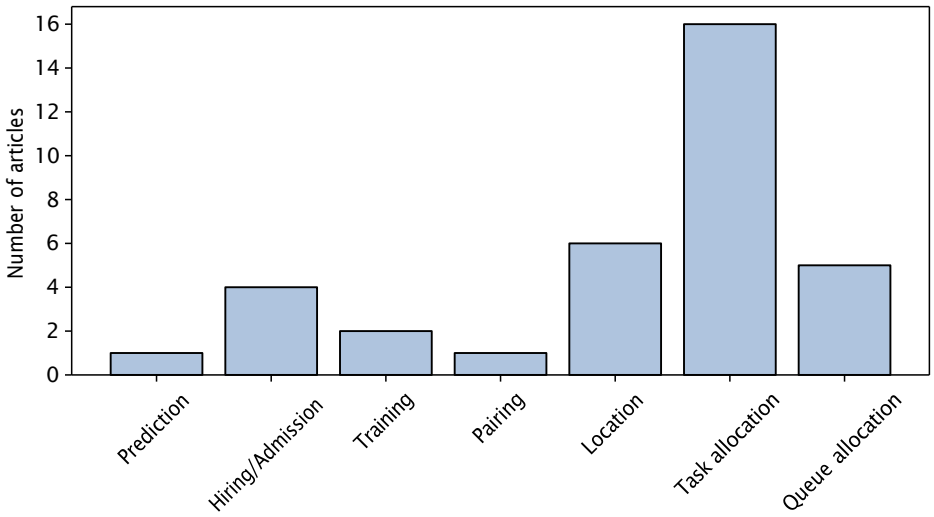


Figure 3. Emergency management decisions

The columns in "Major Decision" identify the key emergency management decisions as introduced above. The columns in section "Objective" specify the modelling objectives, which correspond to the objective functions for mathematical models and the primary evaluation metrics for other methods such as simulations. The column "Preprocessing" indicates whether the authors suggest data transformation actions. These actions can either facilitate data collection (e.g. to derive the demand at a particular location) or be necessary for the model (e.g. to calculate certain metrics). The applied methodologies are detailed in the "Methodology" section. A more extensive review of the papers follows in the next section.

Table 1. Overview - Min: Minimise, Max: Maximise, Opt demand / tasks: Min unfulfilled demand / Max (weighted) fulfilled tasks, Opt satisfaction: Min dissatisfaction (preference, skill, idle time) / Max satisfaction, CTMDP: Continuous time Markov Decision Process

Article	Major decision							Objective										Preprocessing	Methodology							
	Prediction	Hiring/Admission	Training	Pairing	Location	Task allocation	Queue allocation	Opt demand / tasks	Min deviations from demand	Min waiting time	Min reallocation	Max equity	Min required volunteers	Min attrition / rejection costs of volunteers	Opt satisfaction	Min travel cost / distance	Min infeasibility		Linear programming	CTMDP	Agent-based simulation	System Dynamics	Goal Programming	Machine Learning	Heuristics	Other
Falasca, Zobel, and Fetter 2009						x		x							x				x							
Aman et al. 2012 [2]						x			x														x			
Arai et al. 2012					x											x			x							
Falasca and Zobel 2012						x		x							x				x							
Alwahishie and Taaffe 2014						x									x						x					
Lassiter, Alwahishie, et al. 2014			x			x		x					x	x	x				x							
Lassiter, Khademi, et al. 2015			x			x		x									x		x							
Pielorz and Lampert 2016					x								x					x	x							
Mayorga et al. 2017							x	x												x						
Pielorz, Prandtstetter, et al. 2017					x								x					x	x							
Garcia et al. 2018						x		x			x								x							
Rauchecker and Schryen 2018						x		x										x	x							
Sopha and Asih 2018						x		x														x				
Urrea et al. 2019		x		x				x													x					
Abualkhair et al. 2020							x			x											x					
Ozen and Krishnamurthy 2020							x	x												x						
Yang et al. 2020					x											x		x	x							

2Only referring to the humanitarian aid workforce's assignment problem

Table 1 Continued. Overview - Min: Minimise, Max: Maximise, Opt demand / tasks: Min unfulfilled demand / Max (weighted) fulfilled tasks, Opt satisfaction: Min dissatisfaction (preference, skill, idle time) / Max satisfaction, CTMDP: Continuous time Markov Decision Process

Article	Major decision							Objective									Preprocessing	Methodology							
	Prediction	Hiring/Admission	Training	Pairing	Location	Task allocation	Queue allocation	Opt demand / tasks	Min deviations from demand	Min waiting time	Min reallocation	Max equity	Min required volunteers	Min attrition / rejection costs of volunteers	Opt satisfaction	Min travel cost / distance	Min infeasibility		Linear programming	CTMDP	Agent-based simulation	System Dynamics	Goal Programming	Machine Learning	Heuristics	Other
Zayas-Caba´n et al. 2020		x					x			x			x	x						x						
Kebriyaii et al. 2021		x											x						x							
Krstikj et al. 2021					x							x						x							x	
Paret et al. 2021							x			x										x						
S.-Q. Chen, Wang, et al. 2021						x									x			x	x							
S.-Q. Chen, Zhang, et al. 2021						x									x			x	x							
L.-T. Chen and S.-Q. Chen 2022						x									x			x	x							
Moug et al. 2023		x														x			x							
Schmidt and Albert 2022						x		x							x				x							
Sperling and Schryen 2022						x		x				x							x							
S.-Q. Chen and Bai 2023	x																	x						x		
Rabiei et al. 2023						x		x							x			x								x
Zhong et al. 2023					x											x		x	x							
Xue et al. 2024						x									x			x								x

RESULTS

Prediction

To manage volunteers efficiently and avoid a scenario where too many volunteers are deployed and hinder emergency response, emergency managers may first be interested in estimating the number of volunteers required. Prediction models can be useful to support this estimation. S.-Q. Chen and Bai (2023) address this problem and present an approach using machine learning models to predict the damage and derive the number of required volunteers. However, a significant limitation of this model for decision-makers is the focus on typhoons. Additionally, the model was only trained on data from the Philippines, so it is questionable to what extent this kind of model can be applied to a wider range of disasters around the world. Moreover, the authors are not very transparent about their final model parameters, limiting their article's relevance for practitioners. Despite these limitations, the approach described in the article can serve as a guideline for the development of similar prediction models for other types of disasters.

Hiring and Admission

Once the emergency managers can estimate the number of volunteers required, decisions must be made on how many and which volunteers to "hire" or to admit/reject. This problem has been addressed in several papers. At a higher level, Kebriyaii et al. (2021) present a linear optimisation model with the objective of minimising the total costs of the supply chain and casualty management. In addition to the costs of facility location, transportation and inventory, these costs also include the costs of volunteer deployment, e.g. per diem. Although this model does not exclusively focus on volunteers, but rather incorporates them in the general model, it can support emergency managers at a higher decision level.

On a more individual level, Moug et al. (2023) provide a linear mathematical model to decide which volunteers to hire for evacuating people. They propose a two-stage model in which the decision of which volunteer to hire is made before the disaster hits, while the demand for evacuation rides is only known after the disaster has hit. Thereby, the objective is to minimise the costs associated with hiring volunteers and the expected travel costs. The authors present several approaches to solve this rather general model, allowing a practical implementation in a DSS.

Rather than focusing on the cost, but instead on the time required to fulfil "orders", Urrea et al. (2019) present a model to support emergency managers in deciding whether they should allow for congestion in different ranges of supply and demand. Although the authors consider a charity warehouse belonging to a faith-based organisation in their article, we believe that the assumptions regarding volunteers can be translated to SV. Their contribution is twofold, addressing both the question of allowing for congestion and team composition. The presented approach can be useful for emergency managers to determine the optimal number of volunteers per task, especially in spatially restricted areas.

The question of whether SV willing to work should be admitted or rejected upon arrival is addressed by Zayas-Caba´n et al. (2020). The authors analyse a queuing model with the goal of minimising the holding costs for customers, which are assumed to be linearly related to the queue length, the holding costs for volunteers (e.g. costs for equipment and food), and the rejection costs for volunteers. In their model, volunteer managers first have to decide whether a volunteer is admitted. If this is the case, the volunteer is then assigned to one of the queues and can be reassigned at any time. The model is formulated as a continuous-time Markov decision process to solve these questions optimally. While this approach could be implemented in a DSS to provide the optimal answer depending on a given state, i.e. the number of volunteers and "customers" in the system, the authors are aware that in a disaster situation simple policies might be more practical. Therefore, they propose heuristic admission policies to decide whether volunteers should be admitted. While one of the policies only requires the current number of volunteers in the system, the authors also propose more advanced policies that additionally consider the expected workload.

Addressing the problem of hiring and admitting volunteers from very different angles, the presented models can support emergency managers at various levels of decision-making both before and after the disaster has hit. From a cost perspective, the model presented by Kebriyaii et al. (2021) could be integrated into financial decision support tools, as they require several financial metrics as input. In addition to financial metrics, the model presented by Moug et al. (2023) also requires information about the geographic location of potential drivers and evacuees needing a ride. To collect this information, databases of potential volunteers, people requiring help, and geographical information systems can support emergency managers in the decision process. To decide on a level of congestion based on the model presented by Urrea et al. (2019), decision-makers require up-to-date estimates of demand, availability of volunteers and current volunteers in the system, as well as average service times depending on the volunteers' skill level and pairing. Linking this model to sensor data can support the model's accuracy with respect to the real-life representation and thus improve decisions. The same applies to the model presented by Zayas-Caba´n

et al. (2020). Moreover, the heuristic policies proposed by the authors require not only data to tune the model, but also the constant monitoring of certain metrics. In addition to the number of volunteers currently in the system, the more advanced policies further require an estimate of the current workload. To keep track of these numbers, especially with multiple queues and sites, monitoring and tracking systems could assist emergency managers to continuously evaluate these metrics and adjust their policies.

Training

Acknowledging the different skills of volunteers and the possibility of training volunteers, especially when considering longer-lasting tasks such as reconstruction work, represents another potential decision point. In our review, we identified two articles addressing this aspect. In the earlier article, Lassiter, Alwahishie, et al. (2014) propose a linear optimisation model in which volunteers with different skill levels have to be assigned to different tasks requiring different skill levels. While some volunteers already have some skills, emergency managers may also decide to train volunteers to acquire new skills. This, however, limits the efficiency of volunteers during their training, affecting the tasks completed for a particular skill level and therefore the future number of tasks for a particular skill. It also influences the future availability of skilled volunteers and hence their dissatisfaction if these skills are not used. A similar approach is presented in Lassiter, Khademi, et al. (2015). While their model differs slightly in the way of how they deal with skill mismatch, the assumptions regarding training are similar.

Both papers propose a solution for emergency managers to consider the medium-term decision on volunteer training. Since both models need an estimate of the skills required for tasks over a longer horizon as well as an overview of the current skills of volunteers, applications collecting this data could support the usage of these models by providing an up-to-date database.

Teaming

Some tasks in disaster response should be carried out in teams. Surprisingly, little research has been performed on this topic. As introduced above, Urrea et al. (2019) address the question of the optimal team composition in their model for a charity warehouse. The authors differentiate between two levels of experience, "experienced" and "inexperienced", and consider various options of pairing. They show that depending on the level of congestion, demand and available volunteers, different pairing options are preferable to fulfil as many orders as possible.

This model can serve as an initial guideline for emergency managers to decide on how to form teams for specific tasks. However, it is limited to teams of two, which may be applicable for certain tasks such as order-picking but may not be a suitable team size for other tasks.

Location

When volunteers are needed at different sites, emergency managers first face the decision to locate the volunteers. A classic model for this task is presented by Pielorz and Lampert (2016). They adapt a facility location model to minimise the number of volunteers needed, with the restriction that each demand point is covered by a minimum number of volunteers reaching that location within a certain threshold. Although the model is relatively general, it allows emergency managers to analyse the influence of different parameters, such as the required number of volunteers per demand point, on the spatial distribution of volunteers. The authors then extended the case study (Pielorz, Prandtstetter, et al. 2017) and replaced the previously used "flying distance" with a more realistic "driving distance". Although this study does not provide any additions to the mathematical model itself, it gives insights into the differences in locating volunteers when different distance metrics are used.

Krstikj et al. (2021) propose a method to distribute volunteers equitably across neighbourhoods depending on an index reflecting marginalisation. They first suggest an approach to calculate a marginalisation index before presenting a heuristic approach to support the allocation of volunteers to neighbourhoods. Although their case study is based on a pandemic situation and therefore focuses particularly on elderly citizens requiring help, the model can also be adapted to other disaster situations.

A solution approach to match individual volunteers with victims is proposed in the article by Yang et al. (2020). The authors present a heuristic multi-agent reinforcement learning scheduling algorithm to minimise the total time (distance) while ensuring that at least one volunteer rescues each victim. Zhong et al. (2023) propose to use civilian vehicles for evacuation. The authors present a prototype for a constraint-based recommender system that matches volunteers with a suitable vehicle and people who need to be evacuated to a shelter based on constraints on vehicle type and experience, with the objective of minimising the total evacuation time. A similar problem is addressed by Arai et al. (2012). In the suggested approach, volunteers can propose "bids" on disabled persons to help, with the

costs of the bids depending on distance and health status. Using a linear optimisation model, the bids resulting in the minimum overall costs are then selected so that each disabled person is rescued.

The above-mentioned papers propose very different approaches for the spatial allocation of volunteers, ranging from heuristics to linear programming and reinforcement learning. To address the need of up-to-date input data, Yang et al. (2020) suggest analysing data from social media to identify the geographical locations of victims and potential volunteers. However, as regulations regarding access to data from social media platforms potentially change over time, this strategy is not entirely risk-free for emergency managers.

Task Allocation

The vast majority on studies addresses the question of how to allocate volunteers to tasks. One of the objectives in most of these models is either to maximise completed tasks (minimise the uncompleted tasks) or maximise volunteer satisfaction (minimise the dissatisfaction) (Table 1), sometimes in combination with additional objectives. To match the volunteers and the tasks, authors considered different criteria, such as the priorities set by the decision-makers, the volunteers' skills, or their preferences. In the remainder of this subsection, each criterion will be described in more detail.

Collective Level

On a collective level, Sopha and Asih (2018) present a system dynamics model that visualises the impact of allocation policies. In their model, decision-makers must decide on the ratio of volunteers to be assigned to each of the two available tasks, taking into account the dependency between the two tasks. While the model does not support individual decisions, it visualises the overall impact of different assignment policies.

For decision-makers, such a model can be particularly useful when preparing for potential disasters and evaluating different policies. For practical application, future research could therefore focus on linking such a model to databases to retrieve relevant parameters and integrate it into DSS. From an academic perspective, the model could be extended to multiple tasks and their interaction to consider more complex relationships between the tasks.

General Models

Most models aim to support assignments at an individual level. A relatively broad model that addresses this issue is presented by Garcia et al. (2018). The model's objective is to maximise the benefit of allocating resources to tasks (primarily, but not only, volunteers), while also accounting for conflicts and dependencies between different resources. To adapt the model to more specific cases, the authors propose additional constraints, including maximum time limits for volunteers, the allocation of consumable resources, the combination of volunteers with other resources and the inclusion of transport costs.

Rather than maximising the benefit, Falasca, Zobel, and Fetter (2009) present an approach to optimise the allocation of volunteers to tasks by balancing shortage costs due to unfulfilled tasks and deviations from volunteers' preferences and skills with respect to different availability of volunteers and the costs associated with volunteers, e.g. per diem allowances. The model is then extended by Falasca and Zobel (2012) to also account for different group sizes.

Another model that aims to minimise penalty costs caused by a shortage of volunteers as well as deviations from preferences and skills is presented by Aman et al. (2012). Yet, the authors acknowledge that volunteers can be a scarce resource, resulting in an infeasible model. To prevent such a scenario, the authors therefore suggest first solving a model that minimises the volunteers' scarcity before solving the actual model.

These models provide decision-makers with relatively general models for allocating volunteers to tasks, covering the most important aspects. However, even for these models, several parameters need to be known in advance, such as the estimated tasks, the volunteers' preferences, and their skills. In a chaotic disaster situation, it is questionable to what extent these parameters can be estimated or how they could be collected, especially if volunteers arrive continuously. Future research could therefore focus on developing robust practical applications for data collection and optimisation that can be used in disaster situations. To account for the dynamic changes in both tasks and volunteers, another stream of research could focus on developing heuristics and policies that can be applied when tasks and volunteers arrive randomly over time.

Priorities by Decision-Makers

In the aftermath of a disaster, tasks may vary not only in their subjective attractiveness to volunteers, but also in their priority for emergency managers. A model that incorporates these priorities is presented by Rauchecker and Schryen (2018). To ensure that each task is fulfilled at least to a certain extent, the model balances the workload across the tasks by maximising the weighted minimum workload. Moreover, the authors also include constraints to enforce volunteers to work on a single task for a minimum amount of time with respect to their skills and timely availability to avoid chaotic situations and inefficiencies resulting from set-up times.

A strong focus on the different priorities of various tasks is also evident in the model of Sperling and Schryen (2022). Distinguishing between strong and weak priorities, their model aims to assign as many volunteers as necessary to strongly prioritised tasks, while balancing volunteers for weakly prioritised tasks based on quotas. In addition to the time availability and skills, the authors also limit the allocation to non-preemptive schedules.

A model with a special focus on emergency departments is presented by Rabiei et al. (2023). The authors estimate the importance of volunteers' unmet needs and the degree of unsatisfied preferences by proposing a fuzzy interference system. These two metrics are then minimised in a linear mathematical model constrained by the maximum assignment of each volunteer to a single task. To account for the multi-objectivity, the authors present two evolutionary algorithms.

Balancing volunteer allocation between tasks of varying importance can pose a challenge for decision-makers. The models presented above provide different approaches to address this problem. For the practical implementation of these models, several research directions are possible. One potential direction could be the generalisation of these models, as all the authors acknowledge that their model is only tuned for a specific disaster type. Moreover, future research could also put emphasis on the relationship between tasks (e.g. predecessor relationships), or the change in importance over time.

Skills

While most models assume that skills are a fixed parameter, some authors also consider the possibility of training volunteers "on the job". Two models incorporating this option are by Lassiter, Alwahishie, et al. (2014) and Lassiter, Khademi, et al. (2015). As introduced above, both articles present an approach in which volunteers are assigned to tasks according to their skills, but also have the opportunity to acquire new skills through training. While those two models do not really differ in the approach with respect to the training aspect (as discussed above), a significant difference is the mismatch of volunteers' skills and tasks. Lassiter, Alwahishie, et al. (2014) allow for a certain percentage of mismatches, i.e. volunteers are assigned to tasks below their skill level as determined by the emergency manager. On the other hand, Lassiter, Khademi, et al. (2015) do not restrict mismatching by a threshold but assume that the probability of a volunteer leaving depends on the mismatch of skills and thus influences the availability of future volunteers. Moreover, they also associate costs with the mismatch, and include these in the objective function.

To account for the dissatisfaction arising from the mismatch of qualification, Alwahishie and Taaffe (2014) model the problem using agent-based simulation. Their main assumption is that volunteers abandon earlier when dissatisfied, with dissatisfaction arising from both idle time and mismatch of qualification. While the presented agent-based simulation does not suggest an optimal policy, it supports the comparison of different policies with respect to the volunteers' satisfaction and task completion.

By considering the volunteers' skills, decision-makers can increase not only the efficiency, but also the volunteers' satisfaction. While this approach may seem promising in theory, as discussed above, keeping track of the different qualifications and preferences of all volunteers could pose a problem in the aftermath of a disaster. In addition to the research directions already mentioned in the corresponding subsection, future research could also focus on extending the existing models by including bilateral preferences (as introduced in the next paragraphs), e.g. to also consider the preferences of decision-makers for both volunteers performing a task and volunteers being trained.

Bilateral Preferences

Not only volunteers can have preferences regarding tasks, but also decision-makers regarding the volunteers carrying out these tasks. S.-Q. Chen, Wang, et al. (2021) and S.-Q. Chen, Zhang, et al. (2021) address this problem and present a two-sided matching model, in which not only volunteers can place their preferences for different tasks, but also task leaders can indicate their preferences for different volunteer teams. The problem is further extended by considering that volunteers/task leaders may have strong preferences for or aversions to certain tasks/volunteer groups but may not have a preference (positive or negative) for most of the tasks/volunteer groups. To account for the uncertainty that this problem entails, the authors propose an evidential reasoning approach.

The concept of bilateral preferences is also addressed by L.-T. Chen and S.-Q. Chen (2022). In contrast to S.-Q. Chen, Wang, et al. (2021), however, the authors consider individual volunteers (instead of teams) and allow several volunteers to perform several tasks. The optimal allocation is then determined based on a linear optimisation model with the bilateral satisfaction as input.

To collect the preferences of volunteers and their suitability for tasks, emergency managers may use linguistic questionnaires. Xue et al. (2024) therefore propose a method to derive a matrix representing perceived satisfaction. Using this matrix as input, the authors present an algorithm based on foreground theory, differential evolutionary algorithms, and Dempster-Shafer theory to match volunteers to tasks.

The concept of two-sided matching is also included in Schmidt and Albert (2022). However, the suggested approach is unique in that volunteers are offered a list from which they can select a task, rather than assigning them to a task. The authors further extend the model to be also applicable in a multi-epoch setting, and to include individual acceptance probabilities, resulting in customised task recommendations.

For the practical application of the proposed approaches, time might be a relevant factor. Volunteers and decision-makers might be willing to define a few preferences but might be discouraged from filling in long questionnaires. Therefore, future research could focus on predicting preferences based on minimal input from users. Moreover, in addition to preferences for certain tasks, volunteers may also have preferences for collaboration with other volunteers or for certain spatial areas. Including these aspects could increase the volunteers' satisfaction, and thereby, decrease the chance of volunteers abandoning.

Allocation to Queues

Closely related to the allocation of volunteers to tasks is the allocation of volunteers to queues. In these models, queues can either represent counters serving people (disaster-affected people or donors), or assembly lines representing different tasks. The focus of these models is mainly on uncertainty on both sides, on demand, i.e. the arrival of tasks, and on the arrival and abandonment of volunteers. With the exception of the model proposed by Abualkhair et al. (2020), all authors formulate the problem as a continuous time Markov decision process. However, the authors acknowledge the limitations in a chaotic disaster environment, and thus propose practical heuristic policies.

Mayorga et al. (2017) present a problem in which the number of different tasks is fixed and deterministic, while the arrival and abandonment of volunteers and the service rates are stochastic. Each of the different queues is thereby characterised by the number of tasks, the service rate and the maximum number of volunteers that can be allocated to this queue. This problem is then generalised by Paret et al. (2021) to a stochastic arrival of tasks following a Poisson process with rate α_i.

As already introduced above, Zayas-Caba´n et al. (2020) analyse policies for admitting volunteers to a system, and once admitted, assigning, and reassigning them to different queues. These queues differ both in their service time and in their urgency, approximated with holding costs. While several policies are proposed for the decision to admit volunteers, the authors can prove that even for different settings, such as volunteers arriving in batches or working collaboratively on a task, a simply $c\mu$ policy leads to an optimal solution.

The question of how to allocate volunteers to sort donated goods is addressed by Ozen and Krishnamurthy (2020). The authors consider solicited and unsolicited donations, assuming that solicited donations contain only high-priority goods, while unsolicited donation contain a some high-priority goods but also non-priority goods. They aim to support emergency managers in deciding how to allocate a limited number of volunteers to those two queues in order to maximise the throughput of high-priority goods while preventing congestion from unsorted, unsolicited goods.

Another problem related to donations is presented in Abualkhair et al. (2020). The authors model two different queues, one for accepting donations and one for distributing them to beneficiaries. In the scenario where inventories are scarce, the queue of beneficiaries not only depends on the number of volunteers allocated, but also the goods received from donors (and therefore the other queue). The authors then propose a practical solution by suggesting four different heuristic allocation policies.

The models presented can help practitioners to decide on how the allocation of volunteers to queues. In disaster planning, simulating the queues and analysing the consequences of different policies can improve the disaster response. However, this requires a good estimation of parameters such as the arrival or service rates. In the aftermath of a disaster, the practical (heuristic) policies provide a fast decision support. Yet, depending on these policies, different data is required to describe the current state is required. In a chaotic environment, monitoring and tracking systems could support this collection.

FUTURE RESEARCH

As research on the management of volunteers in disaster situations has only been published in recent years, future research can advance all decision problems by extending existing approaches as well as developing new ones.

To the best of our knowledge, there is only one article by S.-Q. Chen and Bai (2023) that addresses the prediction of available volunteers. This leaves room for further research. In addition to extending this model to other disaster types and areas, future research could therefore focus on a more general prediction model that is applicable to multiple disaster types. Another stream of research could shift the focus from the damage to the specific volunteer tasks. S.-Q. Chen and Bai (2023) acknowledged the relationship between the number of volunteers required and the damage caused by the disaster but had to simplify this ratio due to a lack of theoretical background. Future research could therefore focus on the expected number of different tasks depending on the damage and derivation of the number of volunteers required.

There are multiple research directions to further investigate the hiring and admission of volunteers. One potential future direction would be the short-term costs and medium- and long-term effects of rejecting volunteers. While Zayas-Caba´n et al. (2020) consider "rejection costs" in their model to account for the negative consequences of not admitting a volunteer, future research could place more emphasis on the impact of future volunteer availability. In Kebriyaii et al. (2021) and Moug et al. (2023), this effect could be included by also considering the costs if volunteers are not used, while in Urrea et al. (2019) and Zayas-Caba´n et al. (2020), the future arrival rates may depend on the rejection rates. Moreover, future studies could also extend the presented models by including heterogeneous characteristics of volunteers such as experience, skill, or other personal attributes in the decision on how many and which volunteers should be admitted.

A potential future research direction for volunteer training would be the consideration of different, non-hierarchical skills. Volunteers may be highly skilled in one area but lack skills in another. This could result in a volunteer assigned to a task having various levels of satisfaction and efficiency depending on the nature of the task and the skills required. Another future area of research could be the preferences of volunteers in terms of acquiring different skills. In the medium term in particular, the provision of training in the skills favoured by volunteers could have an impact on volunteer satisfaction and therefore on the future availability of these trained volunteers.

Future research on teaming could extend the model by Urrea et al. (2019) to larger team-sizes and more pronounced experience levels. Another potential research direction could also focus on the medium and long-term consequences of pairing. Inexperienced volunteers could benefit from being paired with an experienced volunteer and become an experienced volunteer more quickly than if they were paired with an inexperienced member or vice versa. In the early days, emergency managers may therefore opt for non-optimal pairing strategies in the short-term, allowing them to increase the proportion of experienced volunteers in the medium and long term. For SV, this can be the case a few days or weeks after the disaster.

Several options are available for further research on location planning. First, the model presented by Krstikj et al. (2021) could be extended to more general disaster situations to provide emergency managers with a toolbox for different benchmarks to measure a neighbourhood's marginalisation. Moreover, another research stream could also focus on spatial allocation when tasks differ. Future models could also consider scenarios in which volunteers can easily transfer between locations, but at the costs of time lost for travelling, and scenarios in which this is not possible if the distances between different locations exceeds a certain threshold, thus extending the models presented by Pielorz and Lampert (2016) and Pielorz, Prandtstetter, et al. (2017). To account for the dynamic nature of disasters, future research could apply the approaches by Yang et al. (2020), Zhong et al. (2023) and Arai et al. (2012) to dynamic situations, where the availability of drivers, the need for evacuation and the accessible routes are constantly changing over time. This also raises the question of accurate data collection during a disaster. While some data could be collected in advance, as in Zhong et al. (2023), a relevant part of the data can only be collected once the disaster has hit. To use the proposed approaches efficiently, future research should therefore focus on the collection and integration of this data during a disaster.

With the large variety of models for task allocation, decision-makers have an extensive choice for the combination of different objectives and parameters included. Although most authors approached this question in a relatively similar way using linear programming with the objective of maximising (minimising) the (un)fulfilled tasks or maximising (minimising) the volunteers' (dis)satisfaction (Table 1), the models vary in their focus and emphasis different criteria. While there are still future directions to extend these models from a theoretical perspective, it is questionable whether these extensions would add significant practical value. Therefore, instead of extending existing models, future research could focus on tackling the problems with techniques other than linear programming that account for dynamic environments and also provide solutions for larger instances in a short time.

Some authors have already addressed the managerial implication e.g. by extending their model as in Sperling and Schryen (2022). Answering the question of how to obtain the input data, particularly regarding the volunteers' skills, preferences, and time availability, as presented in Xue et al. (2024), could also simplify the integration of models into real applications. Another approach that could enhance practicality is to give volunteers some degree of choice, as in Schmidt and Albert (2022). To bridge the gap between theory and practice, future research could focus on comparing the usability and practicability of different approaches in disaster situations and proposing "best practises" for future development of DSS.

Regarding the allocation to queues, further research could include different service times based on the volunteers' skills. In particular, if the queues represent specific tasks, as in Mayorga et al. (2017) and Paret et al. (2021), the service time per volunteer may depend on skills and experience (as in Urrea et al. 2019). In addition, the (re)allocation of volunteers to queues may also include set-up times (which can also depend on skills or whether the volunteer is assigned to the queue for the first time). Furthermore, especially when dealing with people as customers as in Zayas-Caba´n et al. (2020), Ozen and Krishnamurthy (2020) or Abualkhair et al. (2020), future research could address the question of fairness between queues with different priorities and thresholds to balance urgency and equity.

CONCLUSION

In this paper we presented a structured literature review of models to support decision-makers in coordinating volunteers in a disaster's aftermath. We identified 31 articles dealing with decisions ranging from predicting the number of volunteers needed to assigning them to tasks and queues. While the literature on task allocation is relatively extensive (Table 1), there is still a lack of research, particularly in the areas of prediction, training, and teams. Moreover, a variety of articles present a deterministic model in which the number, availability, and skills of volunteers and the exact demand is known in advance. This assumption may not hold true in reality, making it difficult for emergency managers to apply these models in the real world. By integrating the two areas of OR/MS and information systems more closely, future research could focus on both the scientific contribution and practical implementation. In particular, the development of versatile models that can be applied to a variety of disaster situations could increase their practical value and facilitate their application.

REFERENCES

Abualkhair, H., Lodree, E., and Davis, L. (2020). "Managing volunteer convergence at disaster relief centers". In: *International Journal of Production Economics* 220.

Altay, N. and Green III, W. G. (Nov. 2006). "OR/MS research in disaster operations management". In: *EUROPEAN JOURNAL OF OPERATIONAL RESEARCH* 175.1, pp. 475–493.

Alwahishie, A. and Taaffe, K. (2014). "Considering volunteer behaviour during relief efforts: A simulation approach". In: Operational Research Society, pp. 182–192.

Aman, A., Bakhtiar, T., Hanum, F., and Supriyo, P. (2012). "OR/MS Applications in Mt. Merapi Disaster Management". In: *Journal of Mathematics and Statistics* 8.2, pp. 264–273.

Arai, K., Sang, T. X., and Uyen, N. T. (July 2012). "Task Allocation Model for Rescue Disabled Persons in Disaster Area with Help of Volunteers". In: *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS* 3.7, pp. 96–101.

Betke, H. (2018). "A volunteer coordination system approach for crisis committees". In: vol. 2018-May. Information Systems for Crisis Response and Management, ISCRAM, pp. 778–785.

Betke, H., Bosse, M., Sackmann, S., and Reusch, C. (2023). "Towards a Taxonomy for Classification of Coordination Systems for Spontaneous Volunteers". In: vol. 2023-text. Information Systems for Crisis Response and Management, ISCRAM, pp. 751–760.

Chen, L.-T. and Chen, S.-Q. (2022). "Volunteer multi-person multi-task optimization dispatch method considering two-sided matching". In: *Soft Computing* 26.8, pp. 3837–3861.

Chen, S.-Q. and Bai, J. (2023). "Data-driven decision-making model for determining the number of volunteers required in typhoon disasters". In: *Journal of Safety Science and Resilience* 4.3, pp. 229–240.

Chen, S.-Q., Wang, Y.-M., Shi, H.-L., and Zhang, X.-X. (2021). "A decision-making method for uncertain matching between volunteer teams and rescue tasks". In: *International Journal of Disaster Risk Reduction* 58.

Chen, S.-Q., Zhang, L., Shi, H.-L., and Wang, Y.-M. (2021). "Two-sided matching model for assigning volunteer teams to relief tasks in the absence of sufficient information". In: *Knowledge-Based Systems* 232.

Daddoust, L., Asgary, A., McBey, K. J., Elliott, S., and Normand, A. (2021). "Spontaneous volunteer coordination during disasters and emergencies: Opportunities, challenges, and risks". In: *International Journal of Disaster Risk Reduction* 65.

Einolf, C. (2018). "Evidence-based volunteer management: a review of the literature". In: *Voluntary Sector Review* 9.2, pp. 153–176.

Fajardo, J. T. B. and Oppus, C. M. (2010). "A mobile disaster management system using the android technology". In: *WSEAS Transactions on Communications* 9.6, pp. 343–353.

Falasca, M. and Zobel, C. (2012). "An optimization model for volunteer assignments in humanitarian organizations". In: *Socio-Economic Planning Sciences* 46.4, pp. 250–260.

Falasca, M., Zobel, C., and Fetter, G. (2009). "An optimization model for humanitarian relief volunteer management". In: Information Systems for Crisis Response and Management, ISCRAM.

Farahani, R. Z., Lotfi, M., Baghaian, A., Ruiz, R., and Rezapour, S. (2020). "Mass casualty management in disaster scene: A systematic review of OR&MS research in humanitarian operations". In: *European Journal of Operational Research* 287.3, pp. 787–819.

Fernandez, L., Barbera, J., and Dorp, J. R. van (2006). "Spontaneous volunteer response to disasters: The benefits and consequences of good intentions". In: *Journal of Emergency Management* 4.5, pp. 57–68.

Fernandez, L. S., Barbera, J. A., and Van Dorp, J. R. (2006). "Strategies for managing volunteers during incident response: A systems approach". In: *Homeland Security Affairs* 2.3.

Fritz Institute (2005). *Recipient perceptions of aid effectiveness: rescue, relief and rehabilitation in tsunami affected Indonesia, India and Sri Lanka*. Fritz Institute.

Galindo, G. and Batta, R. (2013). "Review of recent developments in OR/MS research in disaster operations management". In: *European Journal of Operational Research* 230.2, pp. 201–211.

Garcia, C., Rabadi, G., and Handy, F. (2018). "Dynamic resource allocation and coordination for high-load crisis volunteer management". In: *Journal of Humanitarian Logistics and Supply Chain Management* 8.4, pp. 533–556.

Gutjahr, G. and Viswanath, H. (2020). "A Genetic Algorithm for Post-flood Relief Distribution in Kerala, South India". In: *Advances in Intelligent Systems and Computing* 1077, pp. 125–132.

Kebriyaii, O., Hamzehei, M., and Khalilzadeh, M. (2021). "A disaster relief commodity supply chain network considering emergency relief volunteers: a case study". In: *Journal of Humanitarian Logistics and Supply Chain Management* 11.3, pp. 493–521.

Kovacs, G. and Moshtari, M. (2019). "A roadmap for higher research quality in humanitarian operations: A methodological perspective". In: *European Journal of Operational Research* 276.2, pp. 395–408.

Krstikj, A., Contreras Ruiz Esparza, M., Vargas, J., Escobar, L., Lo´pez de la Rosa, C., Gonza´lez Caldero´n, S., Giorgi, E., and Hinojosa, K. (2021). "Decision-support tool for coordination of volunteers in large-scale lockdowns". In: *International Journal of Disaster Risk Reduction* 62.

Lassiter, K., Alwahishie, A., and Taaffe, K. (2014). "Improving volunteer productivity and retention during humanitarian relief efforts". In: *International Journal of Supply Chain Management* 3.2, pp. 1–10.

Lassiter, K., Khademi, A., and Taaffe, K. (2015). "A robust optimization approach to volunteer management in humanitarian crises". In: *International Journal of Production Economics* 163, pp. 97–111.

Lindner, S., Betke, H., and Sackmann, S. (2017). "Attributes for simulating spontaneous on-site volunteers". In: vol. 2017-May. Information Systems for Crisis Response and Management, ISCRAM, pp. 846–856.

Lindner, S., Kuehnel, S., Betke, H., and Sackmann, S. (2018). "Simulating spontaneous volunteers – A conceptual model". In: vol. 2018-May. Information Systems for Crisis Response and Management, ISCRAM, pp. 159–169.

Mayorga, M., Lodree, E., and Wolczynski, J. (2017). "The optimal assignment of spontaneous volunteers". In: *Journal of the Operational Research Society* 68.9, pp. 1106–1116.

Moug, K., Jia, H., and Shen, S. (2023). "A shared-mobility-based framework for evacuation planning and operations under forecast uncertainty". In: *IISE Transactions* 55.10, pp. 971–984.

Özdamar, L. and Ertem, M. A. (2015). "Models, solutions and enabling technologies in humanitarian logistics". In: *European Journal of Operational Research* 244.1, pp. 55–65.

Ozdemir, U., Mete, S., and Gul, M. (Oct. 2022). "Volunteer Assignment Optimization". In: *Multi-Criteria Decision Analysis*. CRC Press, pp. 31–42.

Ozen, M. and Krishnamurthy, A. (2020). "Resource allocation models for material convergence". In: *International Journal of Production Economics* 228.

Paciarotti, C. and Cesaroni, A. (2020). "Spontaneous volunteerism in disasters, managerial inputs and policy implications from Italian case studies". In: *Safety Science* 122.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., et al. (2021). "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews". In: *Systematic Reviews* 10.1.

Paret, K., Mayorga, M., and Lodree, E. (2021). "Assigning spontaneous volunteers to relief efforts under uncertainty in task demand and volunteer availability". In: *Omega (United Kingdom)* 99.

Patrisina, R., Meilani, D., and Habib, M. (2021). "An information management system for effective disaster relief operations". In: vol. 708. 1. IOP Publishing Ltd.

Pielorz, J. and Lampert, C. (2016). "Optimal geospatial allocation of volunteers for crisis management". In: Institute of Electrical and Electronics Engineers Inc., pp. 152–158.

Pielorz, J., Prandtstetter, M., Straub, M., and Lampert, C. (2017). "Optimal geospatial volunteer allocation needs realistic distances". In: vol. 2018-January. Institute of Electrical and Electronics Engineers Inc., pp. 3760–3763.

Rabiei, P., Arias-Aranda, D., and Stantchev, V. (2023). "Introducing a novel multi-objective optimization model for volunteer assignment in the post-disaster phase: Combining fuzzy inference systems with NSGA-II and NRGA". In: *Expert Systems with Applications* 226.

Rauchecker, G. and Schryen, G. (2018). "Decision support for the optimal coordination of spontaneous volunteers in disaster relief". In: vol. 2018-May. Information Systems for Crisis Response and Management, ISCRAM, pp. 69–82.

Razali, N. A. M., Saad, N. J. A. M., Che Hamid, H. E., Abu Bakar, M. R., Ishak, K. K., Hasbullah, N. A., Zainudin, N. M., Ramli, S., and Wahab, N. (2019). "Volunteer management system for disaster management". In: *International Journal of Recent Technology and Engineering (IJRTE)* 7.5, pp. 569–576.

Salamon, L. M., Sokolowski, S. W., and Haddock, M. A. (2011). "Measuring the economic value of volunteer work globally: Concepts, estimates, and a roadmap to the future". In: *Annals of Public and Cooperative Economics* 82.3, pp. 217–252.

Sampson, S. E. (2006). "Optimization of volunteer labor assignments". In: *Journal of Operations Management* 24.4 SPEC. ISS. Pp. 363–377.

Sauer, L. M., Catlett, C., Tosatto, R., and Kirsch, T. D. (2014). "The utility of and risks associated with the use of spontaneous volunteers in disaster response: A survey". In: *Disaster Medicine and Public Health Preparedness* 8.1, pp. 65–69.

Schmidt, A. and Albert, L. (2022). "Task recommendations for self-assigning spontaneous volunteers". In: *Computers and Industrial Engineering* 163.

Schonbock, J., Raab, M., Altmann, J., Kapsammer, E., Kusel, A., Proll, B., Retschitzegger, W., and Schwinger, W. (2016). "A survey on volunteer management systems". In: vol. 2016-March. IEEE Computer Society, pp. 767–776.

Sopha, B. and Asih, A. (2018). "Human resource allocation for humanitarian organizations: A systemic perspective". In: vol. 154. EDP Sciences.

Sperling, M. and Schryen, G. (2022). "Decision support for disaster relief: Coordinating spontaneous volunteers". In: *European Journal of Operational Research* 299.2, pp. 690–705.

Urrea, G., Pedraza-Martinez, A., and Besiou, M. (2019). "Volunteer Management in Charity Storehouses: Experience, Congestion and Operational Performance". In: *Production and Operations Management* 28.10, pp. 2653–2671.

Van Den Bergh, J., Belie¨n, J., De Bruecker, P., Demeulemeester, E., and De Boeck, L. (2013). "Personnel scheduling: A literature review". In: *European Journal of Operational Research* 226.3, pp. 367–385.

Whittaker, J., McLennan, B., and Handmer, J. (2015). "A review of informal volunteerism in emergencies and disasters: Definition, opportunities and challenges". In: *International Journal of Disaster Risk Reduction* 13, pp. 358–368.

Xue, P., Fei, L., and Ding, W. (2024). "A volunteer allocation optimization model in response to major natural disasters based on improved Dempster–Shafer theory". In: *Expert Systems with Applications* 236.

Yang, Z., Nguyen, L., Zhu, J., Pan, Z., Li, J., and Jin, F. (2020). "Coordinating Disaster Emergency Response with Heuristic Reinforcement Learning". In: Institute of Electrical and Electronics Engineers Inc., pp. 565–572.

Zayas-Caba´n, G., Lodree, E., and Kaufman, D. (2020). "Optimal Control of Parallel Queues for Managing Volunteer Convergence". In: *Production and Operations Management* 29.10, pp. 2268–2288.

Zhong, J., Ngoc, L. L., Negre, E., and Abel, M.-H. (2023). "Ontology-based crisis simulation system for population sheltering management". In: *SIMULATION*, pp. 379–399.

Improving emergency and disaster response management through bidirectional communication between citizens and decision-makers – A field test on the coordination of spontaneous volunteers

Michael Holzhüter
Fraunhofer Institute
for Open Communication Systems
michael.holzhueter@fokus.fraunhofer.de

Hans Betke
Fraunhofer Institute
for Open Communication Systems
hans.betke@fokus.fraunhofer.de

Alina Fröhlich
Fraunhofer Institute
for Open Communication Systems
alina.froehlich@fokus.fraunhofer.de

Michael Klafft
Jade Hochschule & Fraunhofer Institute
for Open Communication Systems
michael.klafft@jade-hs.de

ABSTRACT

Cooperation between different stakeholders plays an important role in managing crises and disasters. The software tools used must also support cooperation and the mutual exchange of information. The bidirectional exchange of information between the population and official agencies poses a particular challenge. In this paper, we present a software demonstrator for the coordination of spontaneous volunteers and its evaluation as part of an overarching research project to gain design knowledge for the development of interconnected systems in disaster management. We describe how the system demonstrator was instantiated on the basis of our existing design knowledge from technical requirements and reference architecture and show the respective interfaces for the stakeholders involved. As the main contribution of the paper, we describe the implementation of a field test to evaluate the demonstrator and present the results in terms of usability and perceived usefulness, which we obtained using a methodologically sound questionnaire.

Keywords

Volunteer Coordination, Field Test, Interconnected Systems, Population, Bidirectional Communication

INTRODUCTION

Good crisis and disaster management depends on how well different stakeholders cooperate (Grolinger et al., 2013). Due to the possible large number of stakeholders involved, there is also a large number of information and communication systems involved. These are usually specialized in storing and processing the data relevant to their user group. To ensure good cooperation, however, it is necessary for the stakeholders to share their information in order to develop a common picture of the situation and determine the best courses of action (Sakurai & Murayama, 2019). This requires systems that enable a digital exchange between stakeholders that is not based on centralized data but is organized in a decentralized manner.

The civilian population plays a special role as a stakeholder in disaster management (DM). On the one hand, they are directly exposed to the risks and consequences of a crisis or disaster scenario and are therefore dependent on comprehensive and reliable information from authorities in order to protect themselves in the best possible way.

(Nikolai et al., 2015). On the other hand, citizens themselves can also be a very extensive source of information to support official DM (Schimak et al., 2015). Bidirectional communication supports the process, as it increases the validity of the information shared, which is essential for making decisions in DM.(Moshtari & Gonçalves, 2017). Kaufhold (2021) conducted a literature review on communication needs from authorities to citizens (A2C) and citizens to authorities (C2A) in crises. The literature review revealed bidirectional communication requirements in crises, thus motivating the need for bidirectional crisis communication tools. A key benefit seen by authorities resp. emergency services workers is better situational awareness through citizen-generated content, such as situational updates, photos, videos or information on the public mood (Reuter & Spielhofer, 2017). Citizens, on the other hand, benefit from information about threats and possible damages, the duration and the area affected by a crisis as well as recommendations about how to cope with the crisis (Reuter, 2014). In general, authorities help to reduce the information deficit of the citizens in times of crises with bidirectional communication solutions (Reuter, 2014).

Studies show that DM authorities and organizations are open to using data from citizens in operations and also to providing information to citizens via bidirectional communication apps (Holzhüter et al., 2023). In the past, therefore, there have already been some research efforts in this area. One specific branch of research that has already been investigated in the past is the coordination of spontaneous or informal volunteers (Whittaker et al., 2015) using bidirectional information systems (Betke et al., 2023; Poblet, 2013). However, although there are already some approaches that address the design and implementation of such systems, e.g. (Batard et al., 2019; Fuchs-Kittowski et al., 2017; Gerstmann et al., 2019), their evaluation is difficult because the systems do not find their way into application and the exercises or experiments required for evaluation are associated with very high effort. As only a few papers have so far focused on evaluating the practical suitability of the approaches, there is a need for further research in this area. In past crises, it has been shown that the use of ad-hoc applications/platforms, such as social networks, e.g. in (Albris, 2018), has enabled spontaneous volunteers to coordinate themselves with all the associated disadvantages, such as unsolicited voluntary actions, data inaccuracies, data overload and even misleading information (Reuter et al., 2015). For better guidance of spontaneous volunteers by DM authorities and organizations for support in coordinated operations, editorial systems are needed that can address end applications, such as volunteer or warn apps. Instead of relying on centralized systems, the use and networking of ad-hoc applications seems promising to harmonize heterogeneous system landscapes and not end at artificial boundaries (federal states, countries, organizations).

In order to extend the knowledge about the practical applicability of such systems, we present the demonstration and field test-based evaluation of a related software artifact. The results of this paper are part of a larger design science research project in which we want to generate design knowledge about the development of interconnected systems in disaster and crisis management. We want to investigate, if the developed architecture, which allows for the use and networking of other ad-hoc applications, is a valid method for providing technologies for various purposes in crises. In the final step of the current research cycle, we have used the design knowledge gathered so far to develop a software demonstrator using the example of a spontaneous volunteer coordination system (SVCS). We chose the example of coordinating spontaneous volunteers (SVs) because an initial comprehensive acceptance study yielded mixed results: without having used such systems themselves, half of the experts were in favor of using such systems, while the other half were either neutral or had a negative opinion (Holzhüter et al., 2023). Taking into account established methods such as UTAUT (Venkatesh et al., 2003), we want to gain insights into the user experience of SVCS, as well as to gain general insights into the suitability of our current design knowledge, in the form of a requirements catalog and a reference architecture, for the development of practicable software. The collected feedback serves as a foundation for targeted improvements aimed at increasing acceptance in crisis situations. Our current efforts are directed towards ensuring that the systems not only provide utility and user-friendliness but also align with user expectations. With the results presented here, we want to facilitate the future development of interconnected systems and especially SVCS for scientists and practitioners, as well as to show a possibility for the methodical evaluation of corresponding systems with regard to perceived usefulness and usability. For this purpose, we want to examine two research questions in this paper:

(Research question I) *To what extent is our existing design knowledge suitable for the development of application-ready systems?*

(Research question II) *How do the two user groups disaster managers and spontaneous volunteers evaluate a special coordination software in terms of usability and usefulness under practice-oriented conditions*?

This paper is structured as follows: In Section 2, we explain our approach based on a design science research method, and then in Section 3, we describe our instantiated software demonstrator based on our previous gathered design knowledge. In section 4 we describe the experimental setup of our field test to evaluate the software demonstrator In Section 5, we present and discuss our evaluation results. Section 6 contains a brief discussion of related work, particularly in relation to the evaluation of similar software artifacts. Finally, in Section 7 we give an outlook on further research so that our proposal for a reference architecture can be implemented in practice.

RESEARCH METHODOLOGY

It is our superordinate objective to create a theorical IS artifact that helps developers, disaster response experts and researchers to create interconnected systems for DM. Designing such an artifact is a demanding task, requiring innovative approaches, particularly in domains where existing theories fall short (Hevner et al., 2004). To accomplish this goal, we have initiated a design science research (DSR) project, following Hevner et al. (2004). DSR, as defined by Iivari and Venable (2009), involves "[…] a research activity that invents or builds new, innovative artifacts for solving problems or achieving improvements […]" (Iivari & Venable, 2009). By combining scientific principles with practical relevance, we aim to provide valuable design knowledge for the development of ICT in DM (Hevner et al., 2004). Our approach, therefore is based on the design science research methodology by Peffers et al. (2007). In this section we describe how we instantiated the method of Peffers et al. (2007) as our IS research framework to structure the iterative process to design and evaluate a suitable IS addressing our research question (see Figure 1). By employing this framework, we ensure the incorporation of continuous improvement throughout the artifacts development process. As our framework starts with the identification of a problem and motivation, we follow Peffers et al.'s (2007) problem-centered approach to guide our investigation, driven by both practical and scientific demand.

	Identify Problem & Motivate	Define Objectives of a Solution	Design & Development	Demonstration	Evaluation
Instantiation of Research Steps following Peffers et al. 2007	**Method:** Survey on the need for and acceptance of information and communication systems for the exchange between disaster response officers (n = 354) and the public (n = 504). **Result:** Both user groups have a fundamental need and willingness for information exchange in different use cases using an IT system, including a system for coordinating spontaneous voluntary tasks. *(Holzhüter et al., 2023)*	**Method:** Literature analysis and expert survey (n = 49) to derive requirements for a decentralized multi-stakeholder information and communication system in disaster management **Result**: 40 requirements identified (38 functional, 2 non-functional) *(Holzhüter & Meissen, 2021)*	**Method:** Conceptual description of a solution based on the derived requirements **Result**: Service-oriented reference architecture for a decentralized multi-stakeholder information and communication system in disaster management *(Holzhüter & Meissen, 2021)* --- **Method:** Implementation of a software artifact based on the reference architecture using the example of a coordination system for spontaneous volunteers **Result:** Web-based dashboards with widgets for the integration of various situation information and a module for coordinating spontaneous volunteer work. iOS & Android app for registration, alerting and coordination of spontaneous volunteers	**Method:** Application of the software artifact (dashboard & app) in a realistic use case with both user groups **Result**: Planning and execution of a disaster response exercise for the reception of refugees with the participation of disaster management employees (n = 13) and spontaneous volunteers (n =32)	**Method:** Survey among the use case participants to evaluate the usefulness and usability of the software artifact using UTAUT-methodology (Venkatesh et al.(2003)) **Result**: Evaluation results for 22 items per user group plus comments from participants show a consistently positive perception of the software artifact.

Figure 1. Design science research methodology process model by Peffers et al. (2007)

As part of the first step, **problem identification and motivation**, we conducted a broad-based survey among the public (n = 504) and DM professionals (n = 354) to identify specific challenges in the area of information systems for information exchange. In summary, the results showed that participants from both DM and the public were overwhelmingly positive about and in favor of information exchange in various use cases (Holzhüter et al., 2023). The survey showed that there is a need on both sides to provide and receive different types of information. Interconnected systems are a promising approach for bringing together a wealth of information from different sources and processing it in a user-friendly way.

To **define the objectives of a solution**, we drew up a catalog of requirements. This is based on an extensive literature analysis and a survey of experts in the field of interconnected systems development and experience in DM (n=49). As the focus of the survey was on functional requirements, we were able to identify significantly more, resulting in 38 functional and 2 non-functional requirements (Holzhüter & Meissen, 2020). The concrete identification and consideration of the undoubtedly existing non-functional requirements represent a good approach for the processing of a further research cycle.

The **design & development** of a suitable solution consists of a general theoretical part and a use case-related

software development. We have designed and described a reference architecture as a general theoretical basis for the implementation of interconnected systems in DM (Holzhüter & Meissen, 2020). This shows how stakeholders with different applications and technical requirements can exchange information with the help of a decentralized information exchange adapter.

Up to this point the results were generated in previous papers, all following steps are the core contributions of this paper. To further concretize the solution approach, we have developed a software demonstrator considering the requirements and reference architecture. As our acceptance study showed that users are generally positive about a SV coordination system, but that there is also a lot of skepticism, we decided to develop a software demonstrator in this problem area to investigate user acceptance of a specific system. The resulting demonstrator is SVCS, consisting of an iOS & Android app for the registration and alerting of SVs, as well as a web-based dashboard for the creation and monitoring of SV missions. The system is described further in the following chapter.

For the **demonstration**, we carried out a disaster response exercise in a medium-sized German city (approx. 99,500 inhabitants) with the participation of SVs (n = 32) and local DM staff (n = 12). In the exercise context of a mass arrival of refugees, both groups of participants used their respective subsystems of the VCS to manage and process multiple missions. The exercise demonstrated that the processes and tasks in the application area can be supported by an appropriate system.

The **evaluation** was carried out in connection with the demonstration. The exercise participants were asked afterwards to complete a questionnaire on their experience of using and perceived usefulness of the VCS used. The questionnaire was methodologically based on the UTAUT model by Venkatesh et al. (2003). UTAUT, as one of the most common methods for evaluating software artifacts, takes into account other approaches such as the Technology Acceptance Model (TAM), the Theory of Reasoned Action (TRA) and the Theory of Planned Behaviour (TDB). Hence, it offers a comprehensive perspective on the various factors that influence the acceptance and use of technology. The items surveyed by UTAUT are familiar to many scientists and increase the comprehensibility and comparability of the evaluation results in this paper. The results showed a high acceptance of both user groups when using the system. A detailed description of the demonstration and evaluation results can be found in sections 4 & 5.

INSTANTIATION OF THE SOFTWARE DEMONSTRATOR

The objective of the solution is derived from a reference architecture, which was created on the basis of a requirements analysis (Holzhüter & Meissen, 2020). The requirements were derived from a user survey and existing solutions (Holzhüter et al., 2023) in a larger acceptance study. For this purpose, various use cases were created that can better involve the population in disasters. The use cases include the feedback of warning messages (work in progress), the inclusion of data from the population, e.g. private weather stations (Holzhüter et al., 2021) the coordination of SVs (evaluation in this paper) and a feedback function for the population in the warning process (not part of the paper).

The demonstrator focused mainly on interfaces and, for the demonstration in the research project, on visualization using dashboards and widgets for data visualization. In addition, a volunteer component was added to the KatRetter app and a feedback function was added to the KatWarn app as value-added services. The KatRetter app, with the actual aim of coordinating first responders for cardiological emergencies, was extended by an independent KatHelfer component. The KatHelfer component is a value-added service that takes over the coordination of SVs. The following requirements in Table 1 from the reference architecture (Holzhüter & Meissen, 2020) were implemented as described.

Table 1. Mapping of requirements to developed services

ID	Requirement	Description	Realization in the demonstrator
R01	Authentication	Authentication should be state-of-the-art and support multiple authentication methods.	State of the art Authentification-Service implemented.
R02	Multi-client capability	The system must be multi-client capable. It must be possible to configure role-based data display and transfer.	A roles and rights model has been created. This allows different roles to have different views of widgets, for example to create assignments or approve assignments as a decision-maker.

ID	Requirement	Description	Realization in the demonstrator
R03	Filter option	Filter options should be offered; the focus should be on the relevant data.	Data can be freely filtered. When configuring the widgets, filters can be applied to certain regions, certain radiuses, certain deviation criteria or even certain alerts, for example, to create the optimum focus for the user.
R04	Validation of information	A transaction check is to take place. The data must be unchangeable and usable in court.	"Deviation criteria" have been developed for various widgets so that "incorrect" data does not distort calculations. Correct information is to be validated using a blockchain, which is still in the concept phase.
R05	Data exchange	Data should be usable not only in the event of a crisis but also during normal operation to generate synergy effects.	Some data is stored in a local database as an OLAP cube to efficiently access data outside of events, for example for follow-up. To improve data exchange, care was taken to ensure data anonymity when designing the interfaces so that no additional data protection requirements had to be implemented.
R06	Event orientation	It should also be possible to view the data under less-than-optimal conditions, for example when information systems have failed. The system must act event oriented.	The high level of interface orientation means that other connected systems can continue to operate in the event of system failures. In addition, work can continue the existing "old data" until the interface is available again. The user is visually informed when data falls out of the update rhythm.
R07	Data actuality	The system should be able to update information automatically.	The interval at which data should be updated can be set in the dashboard.
R08	System boundary	The system should also be applicable in federal states.	The system design knows no municipal or national borders.
R09	Automation	The system should be fully automated and should not require manual operations.	No further input is required for widgets created by the user.
R10	Prioritization of data transmission	The system must offer the possibility to prioritize certain transactions to certain participants.	Implemented by an ETL service.
R11	Geoinformation system	The system must offer the possibility to transfer data from geographic information systems.	Map service is implemented and can be used as a form of visualization for most widgets.
R12	Interfaces	Standards like OGC (Open Geospatial Consortium), STANAG (NATO contracting parties' Standardization Agreement) should be usable. The exchange format should be easy to design. The interface should be compatible with as many media as possible in terms of hardware and software. The data network should be uniform, and services usable across systems. It should also be possible to establish permanent connections to check availability.	Standardized interfaces were implemented. The requirements have been fully implemented. Interfaces can also be used by other value-added services so that the application can be used not only with the KatRetter or KatWarn apps. However, the app developers are responsible for implementing the interfaces on their side.
R13	Control over connections	The system must be able to give the user full control over the data connections.	Is controlled via the rights and roles system.

ID	Requirement	Description	Realization in the demonstrator
R14	Use of own software, transformation of data	The system must be able to transform data so that users can use their own systems.	Not implemented in the use case. In this case, the company's own dashboard was developed. As part of SPELL, data is also to be made usable in third-party software via interfaces.
R15	Data visualization	The system must be able to prepare data in such a way that it can be displayed within the system if the data is not transferred to another system by the user.	Dashboard with widgets was implemented for this use case.
R16	Central storage of information	The system must also be able to store data centrally so that the user can analyze them even without his own software.	Backend of the dashboard holds the data in OLAP cubes.
R17	Trusted connections	The system must be able to easily establish communication connections to other users.	This use case has been implemented. Additional portals from private weather stations can be easily integrated, as well as additional value-added services such as apps for warning messages or the coordination of SVs.
R18	Digital data transfer	The system must be able to support many different formats and interfaces so that data can be transmitted digitally.	This was the focus of the implementation, so that there were no media disruptions.
R19	Real-time transmission	The system must be able to transmit data in near real-time (e.g. video streams)	This was not necessary for the application.
R20	Filter options	See R03	See R03
R21	Interfaces	The system must be able to create, adapt or delete interfaces	See R12
R22	Decentralized system architecture	The system must be decentralized to increase the availability of information.	The system is designed to be decentralized. The aim is for each user to use their own system, as shown in Figure 2, or to use the dashboard with their own local storage. This means that the architecture is not dependent on any central data or systems.
R23	Open-Source	The system must be open source so that other providers can make extensions (including interfaces, value-added services) quickly available. In addition, a reference architecture should make it possible for different service providers to develop the system.	During system development, care was taken to ensure that free libraries were used.
R24	Validation of information	See R04	See R04

From the continuous evaluation of user feedback and results from previous surveys, on which the requirements analysis is based, and a consideration of the advantages and disadvantages of various architectural models, the following non-functional requirements emerged. When introducing applications, not only the non-functional requirements of end users (such as reliability, user-friendliness, simplicity, intuitive operations, understandability) must be satisfied, but also those of administrators (such as scalability, traceability, interfaces, innovative capability, maintenance) or management (such as it-security, data security and privacy, auditability, costs). These were also evaluated by stakeholders as part of the reference architecture (Holzhüter & Meissen, 2020).

Figure 2 shows the reference architecture from (Holzhüter & Meissen, 2020). The grayed-out areas are the components that were not implemented. For the demonstrator, the part in which the data was processed via the web browser was implemented. In addition, the reference architecture was extended by the stakeholders for value-added services.

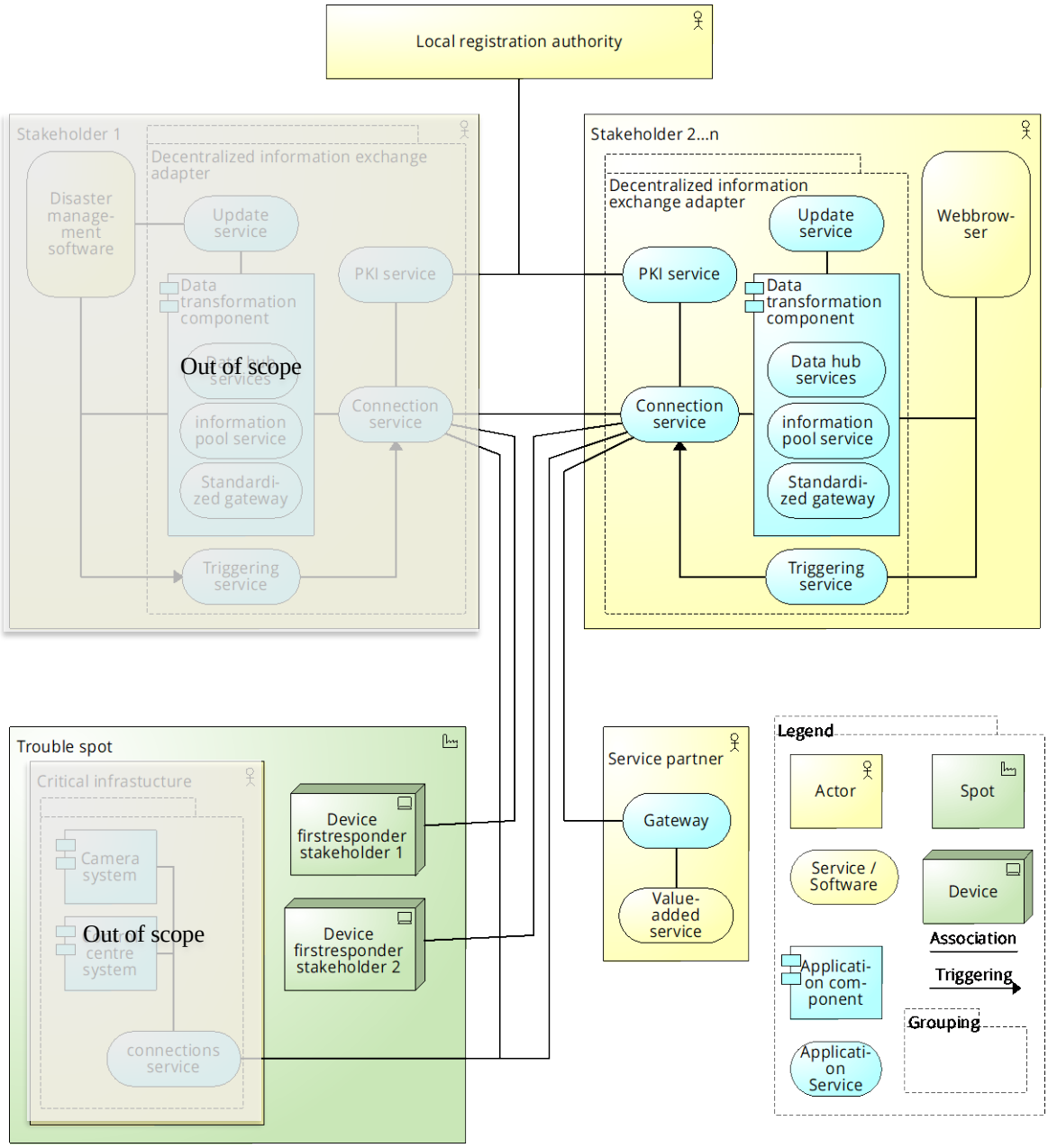


Figure 2. Adapter for decentralized information exchange with value-added services

The components for displaying information were developed in the form of an easy-to-configure dashboard with widgets (see Figure 3). On the left side, new dashboards can be created to have different views for various situations, for example. Each user can create their own dashboards and customize the widgets in the grid differently. This allows most widgets to be displayed as statistics, key figures, maps, or even heatmaps.

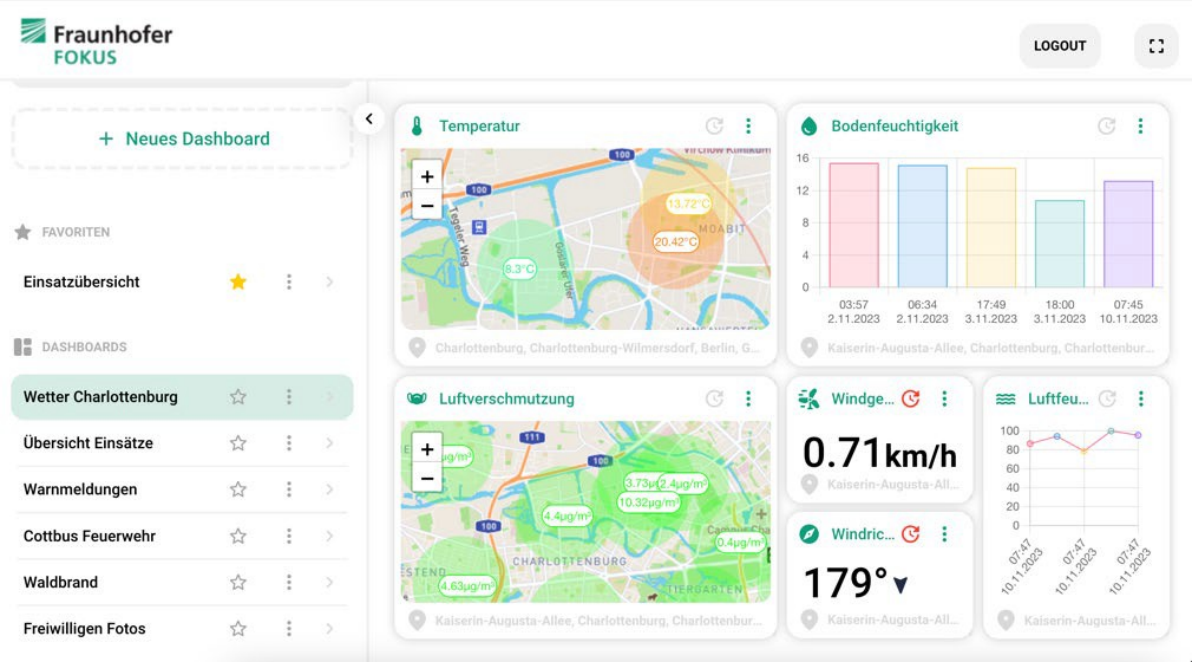


Figure 3. Example of a configured dashboard with widgets for displaying weather data from private weather stations

Figure 4 shows the implementation for displaying the evaluation of warning messages (not the use case of this paper). This shows that the demonstrator also works for other applications of bidirectional communication. In this scenario, the user is presented with information pertaining to an alert notification within a specified area. The user receives an overview, including the number of users who received the alert in the designated area, the quantity of individuals who provided feedback on the alert, and the timeframe during which the feedback was received. This feature will not be discussed further because it was not part of the exercise.

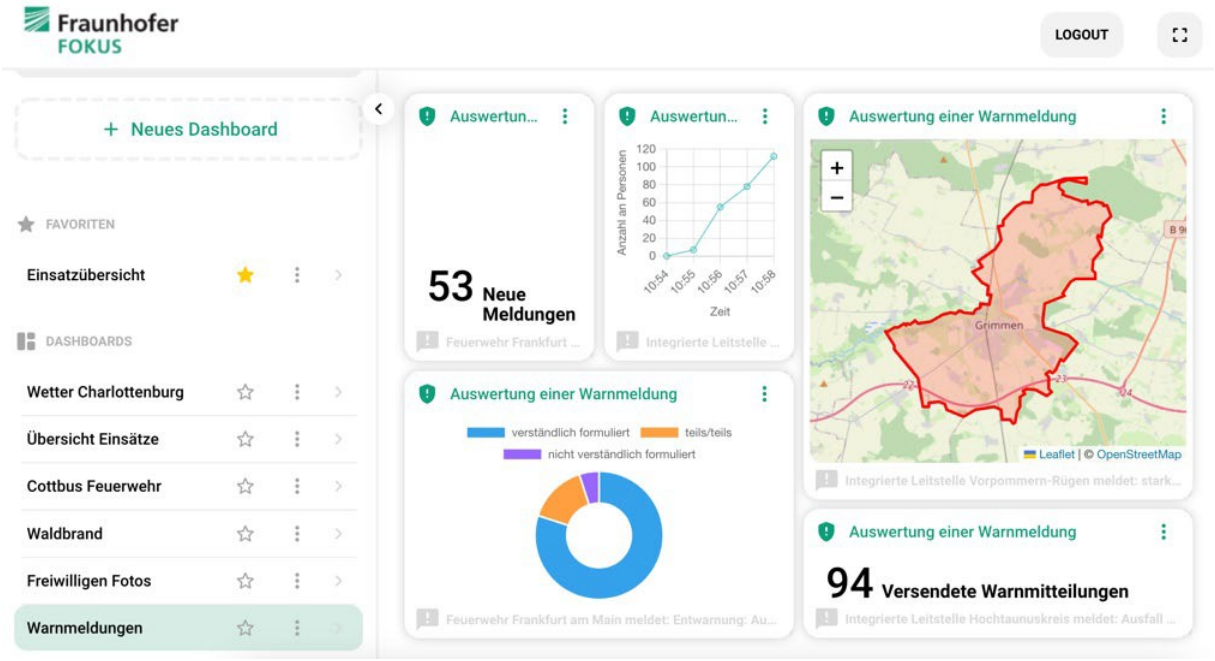


Figure 4. Example of a configured dashboard with widgets for evaluating feedback from alerts

Value-added services, as used in this exercise, represent additional services that cannot be mapped in the adapter. The KatHelfer component designates/ allocates tasks to the right registered volunteers. The component can use the KatRetter app for this, but other services for SVs can also be used. The adapter for decentralized information exchange handles communication with the KatHelfer component. The dashboard is used to create deployments. Figure 5 shows the option of creating assignments for SVs in the dashboard. Specify when the operation starts and ends, what type of operation it is (this becomes crucial for the app settings in Figure 7), what the operation

should be named, provide a description of the operation, indicate the number of SVs needed, specify the meeting point, and identify the geographical area in which the operation is located.

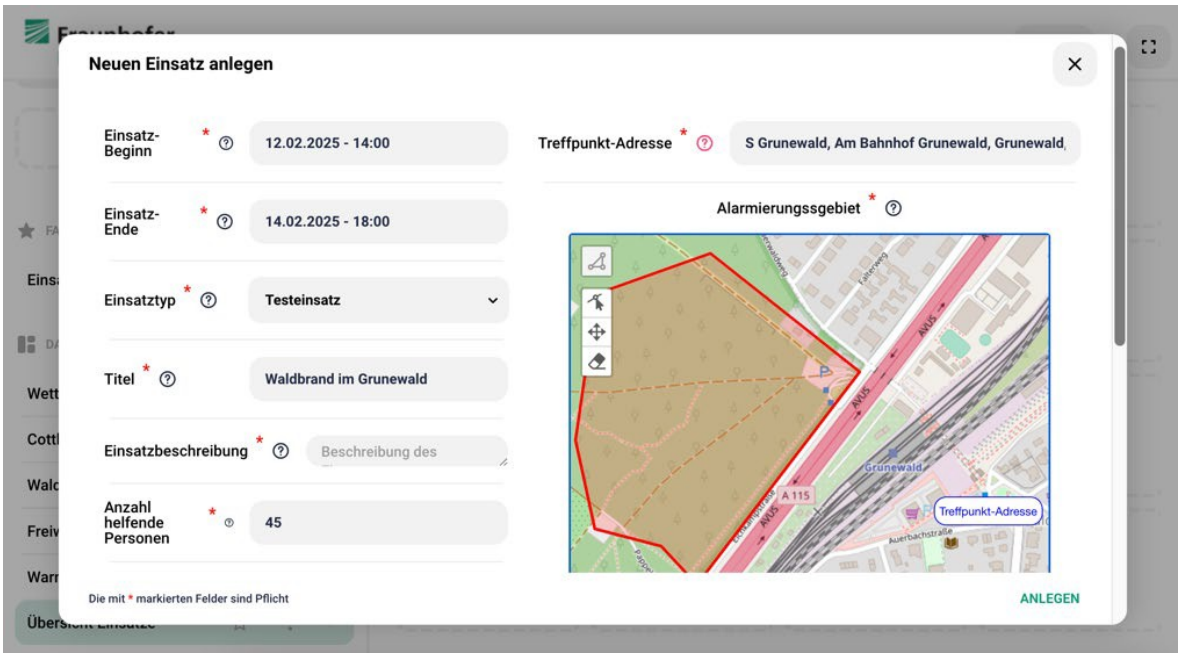


Figure 5. Example of creating an operation in the dashboard

Once the assignment has been created in the dashboard, the assignments are displayed in an overview (see Figure 6). A description, the status and the number of activated volunteers is displayed. The assignment can also be ended via a shortcut.

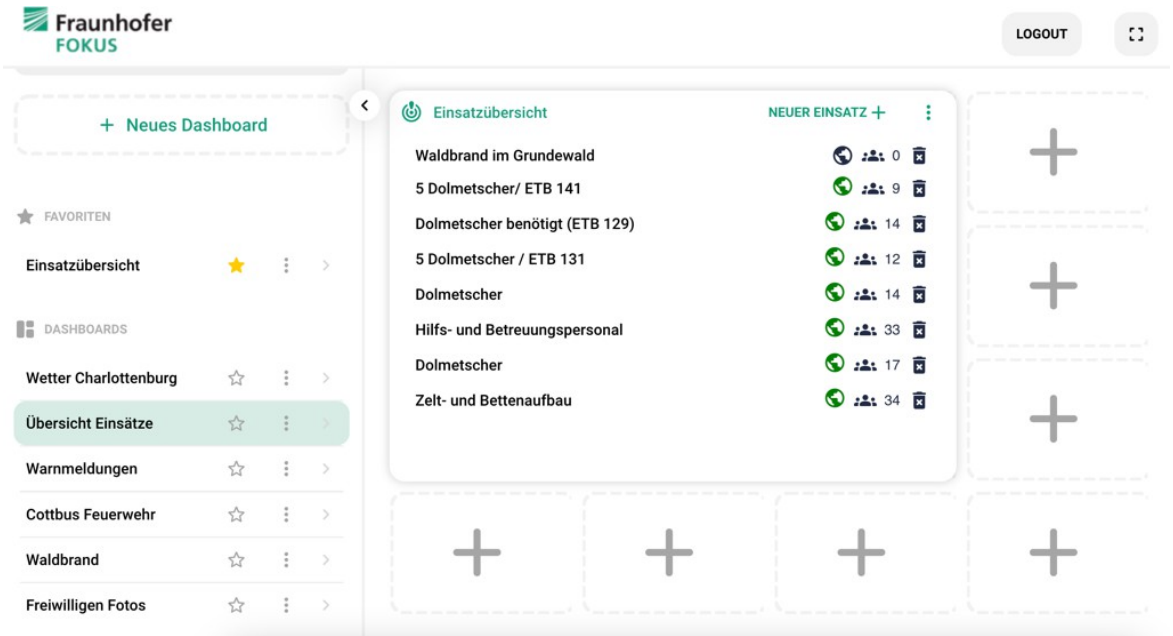


Figure 6. Overview of operations

Value-added services, as used in this exercise, represent additional services that cannot be mapped in the adapter. The adapter for decentralized information exchange handles communication with the KatHelfer component. The dashboard is used to create missions.

To utilize the KatHelfer component within the app, users must configure certain settings, as illustrated in Figure 7. This enables users to specify their availability as a volunteer (named “Mithelfer” in the picture on the left) during disasters. Subsequently, users can select the types of missions for which they wish to receive notifications (in the picture on the right), such as those related to catering support or missions requiring more physical

involvement.

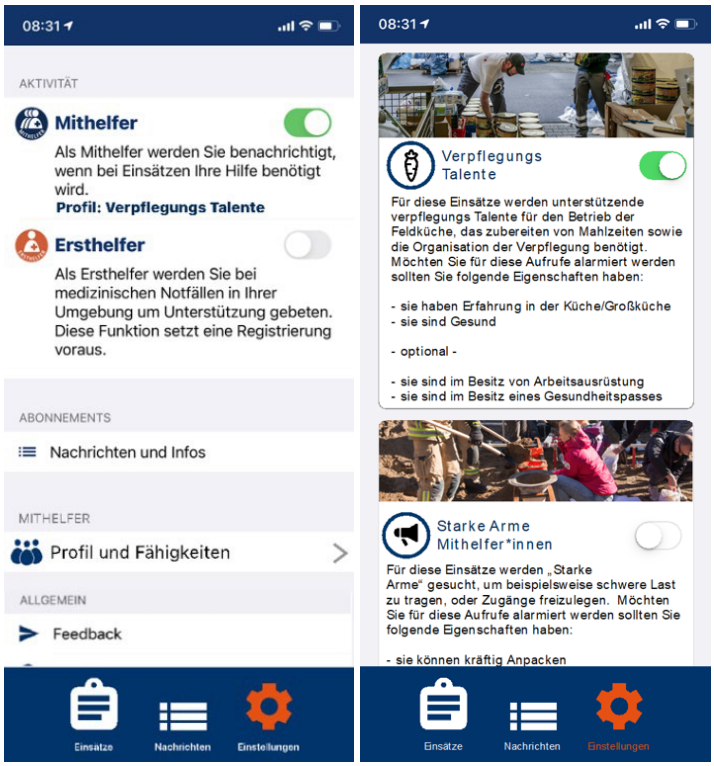


Figure 7. Settings

If volunteer assistance is needed, an operation is created and sent to volunteers, as depicted in the dashboard above. Volunteers can now view the operation details, as shown in Figure 8 (on the left side). Users can review the operation details and decide whether they wish to support the operation (as shown in the middle image). If they opt to do so, they can proceed to the operation site, confirm their on-site presence, or choose to cancel their participation in the operation (as shown in the image on the right).

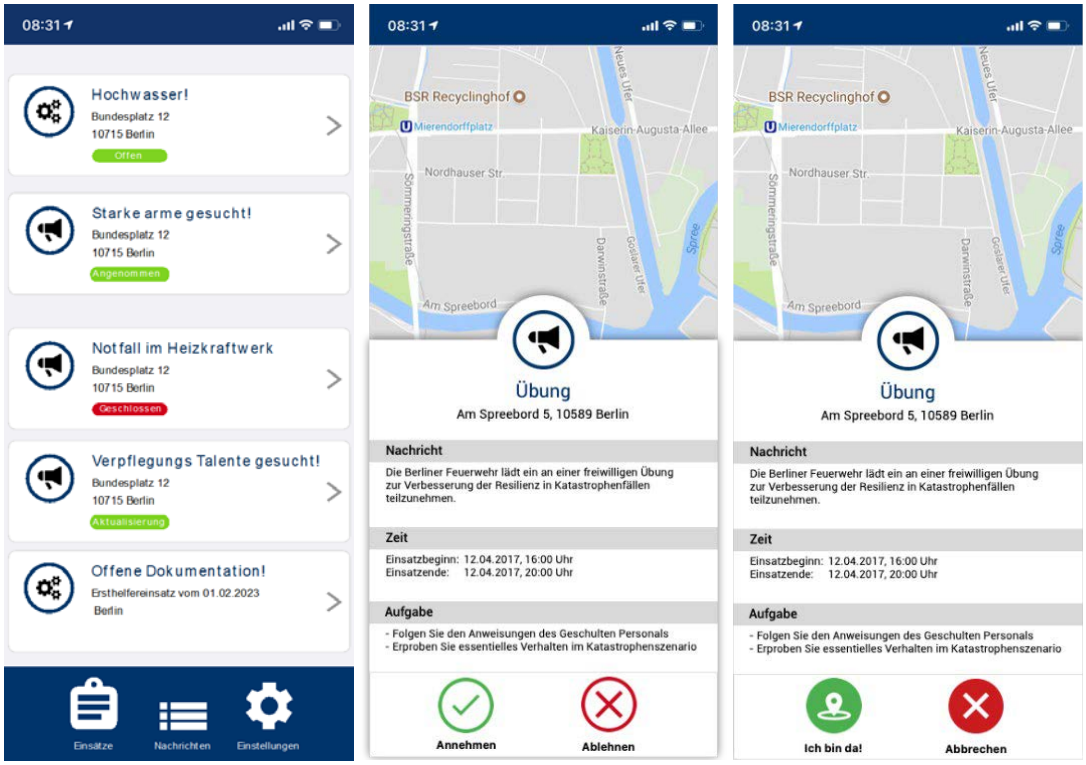


Figure 8. Mission Screens (Mission List, Accept/Decline Mission, Arrived/Mission Aborted).

The acceptance study (Holzhüter et al., 2023) revealed that there is a desire for technological support for spontaneous interventions by voluntary volunteers from both the general population and emergency response forces. However, a crucial aspect is the need for coordination of SVs. To address this, a tool has now been provided to both stakeholders, which needs to be evaluated in use before it can be employed in actual operations.

With the development and presentation of the SVCS in this chapter, we can make initial statements regarding our first research question: We have shown that it is possible to instantiate a software based on our current design knowledge using a concrete application domain. We have shown how the technical requirements can be addressed and how the reference architecture is used, thus providing starting points for the instantiation of further systems. In the following chapter, we show how we demonstrate the suitability of the software for the application and how we have set up the evaluation to answer the research question II.

FIELD TEST & QUESTIONNAIRE DESIGN

For demonstration and evaluation of our solution we conducted a disaster response exercise with both user groups. The exercise was set in Cottbus, a middle-sized city (99.500 residents) in Germany. Aim of the exercise was testing the capabilities of the city's rapid intervention unit and evaluation of the KatHelfer component. A new organizational aspect of this exercise was the involvement of SVs. The exercise was based on a scenario where a large group of refugees arrived from Ukraine. The DM was expected to set up care facilities for the refugees, provide them with medical care and food, register them and look after them until their onward journey, as well as coordinate that journey. Thus, various tasks needed to be completed in order to manage the situation. These tasks included setting up tents and camp beds, providing language mediation for Russian, Ukrainian, Polish and English at various locations, setting up a supportive IT structure and offering nursing and non-acute medical care for the refugees. The tasks were to be accomplished with the help of SVs from the civil population. Throughout the course of the exercise, it was therefore necessary to register, activate and coordinate SVs via the KatHelfer component to manage the situation.

The members of the rapid invention unit served as the coordinating disaster managers (n = 13) while the SVs were people who had already been active as volunteers in previous disaster situations (n = 32). The disaster managers had an average age of 36 years, with a range of 17-61 years. 5 identified as male, 3 as female, 5 did not specify their gender. The majority had a German high school diploma or a comparable qualification. The SVs were on average 44 years old; the youngest participant was 20 years old and the oldest 65 years old. 14 of them were male, 8 female, 10 did not specify. 61% of the volunteers have a German high school diploma or comparable. The exercise began at 9 a.m. with the alert that 600 refugees, in various states of health, would arrive at 10 a.m. in Cottbus and ended at 12 p.m., with the dismantling and the announcement via the KatHelfer component that the exercise was now over. At 12:45 p.m., the units were then moved back to their original locations. Afterwards, the exercise was evaluated with the help of our questionnaire. In addition, short interviews were conducted with some of those involved and observations were made and noted during the exercise, both in DM and on site.

The aim of the evaluation was to gain insights into the perceived usefulness and user-friendliness of the KatHelfer system demonstrator for both user groups (SVs, disaster managers). For this reason, we based the questionnaire on items and constructs from the established UTAUT (Unified Theory of Acceptance and Use of Technology) (Venkatesh et al., 2003). One problem with using the UTAUT, however, is that not all items fit the context of our specific application. For example, the item "If I use the system, I will increase my chances of getting a raise" was not applicable to our scenario since the dashboard and app are not intended to provide an economical benefit for the users. For the same reason of transferability to our specific scenario the item groups that measure the factors "social influence" and "facilitating conditions" were not used, as well as the two items, "It scares me to think that I could lose a lot of information using the system by hitting the wrong key" and "I could complete a job task using the system if I had just the built-in help facility for assistance". In the end, we were left with 14 items of the UTAUT, measuring Performance Expectancy, Attitude towards using technology, effort expectancy, self-efficacy, anxiety, and behavioral intent. To gain further information on how our design is perceived and about user satisfaction, we additionally added 8 items from a questionnaire from (Sindhuja & Dastidar, 2009), which is derived by the Purdue Usability Testing Questionnaire (PUTQ) by (Lin et al., 1997). In conclusion our questionnaire contained 22 items, measuring user acceptance of our dashboard and app with main emphasis on usefulness and ease of use/design. We translated these items to German and developed two versions, one for the members of the rapid intervention unit to rate the dashboard and one for the SVs to rate their experiences with the app. In addition, we added demographical items about age, gender, level of education and profession, taken from the recommendation for demographical standards from the German Federal office of Statistics.

EVALUATION RESULTS

Of the 13 DM experts using the dashboard, 10 completed the questionnaire in full, we consider. Even if this

number of participants is too small for a comprehensive statistical evaluation, this number of experts already provides useful results regarding the usability of software artifacts, as described by (Hwang & Salvendy, 2010) in their 10 +-2 rule. Of the SVs, all 32 participants completed the questionnaire. Of the SVs, all 32 participants completed the questionnaire. This difference in the number of participants should be considered when examining the results presented below as a percentage. The results are presented and briefly summarized in the following based on the dimensions surveyed and the associated items for both user groups responsible frontends (dashboard/ app).

Performance Expectancy: Both user groups rate the usefulness of the system used as predominantly positive (Figure 9). Usefulness in work as SV is rated considerably higher for the app than for the dashboard, going in line with the results of the acceptance study in our earlier research steps.

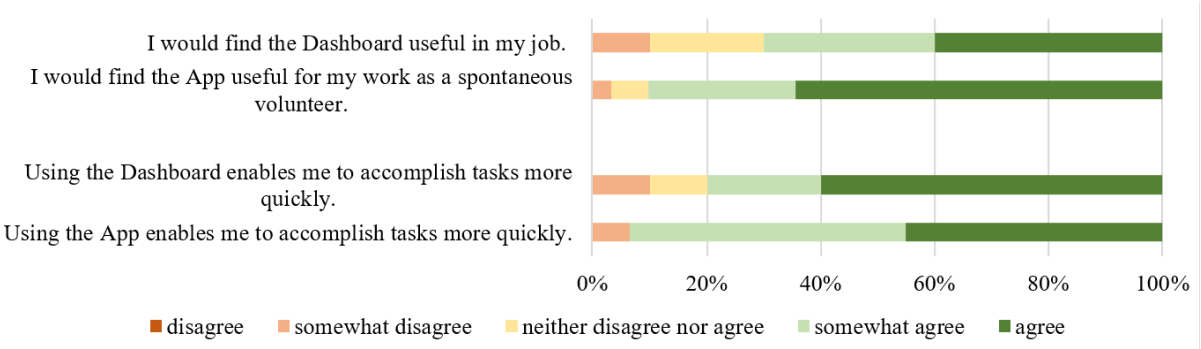


Figure 9. Performance Expectancy

Attitude Towards Technology: The results of this dimension show that both user groups enjoyed interacting with the system by a large majority (Figure 10). The negatively polarized item generates responses with a wider range of agreement to disagreement for the app. However, this pattern is repeated for the other negatively polarized items from other dimensions. Perhaps a percentage of SVs also missed the polarity reversal (however, this does not explain all "Agree" and "Somewhat Agree" responses, because the percentage is lower for the other reversed items). The results of the reversed items show that, overall, the items were read and answered carefully.

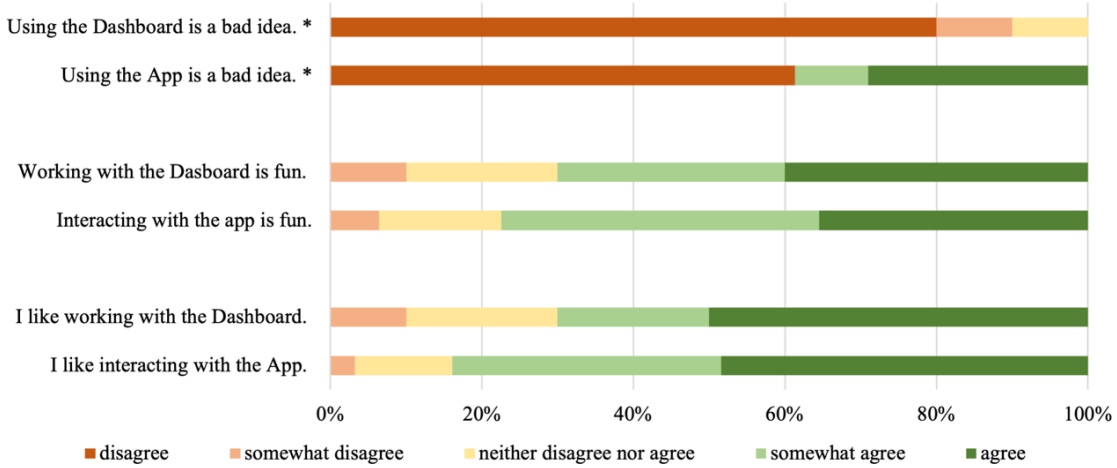


Figure 10. Attitude Towards Using the Technology (* negative polarized items)

Anxiety: The results show that both user groups have few reservations about using the system (Figure 11). However, it is evident that the SVs are more skeptical about using the app. One possible explanation could be that some of the SVs who have so far coordinated via existing communication channels see no need to introduce further systems and are fundamentally skeptical about unknown applications.

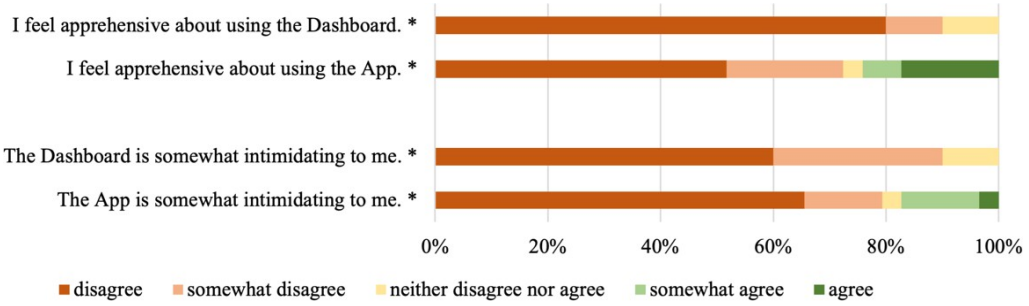


Figure 11. Anxiety (* negative polarized items)

Behavioral Intent: While most respondents in both user groups also had a positive attitude towards the system in the case of behavioral intent, a slightly higher level of rejection was found among SVs (Figure 12). However, the fact that some completely reject the use of the system is consistent with the results of some other dimensions (e.g. satisfaction, anxiety), which show a clearly negative attitude among a small subset of SVs.

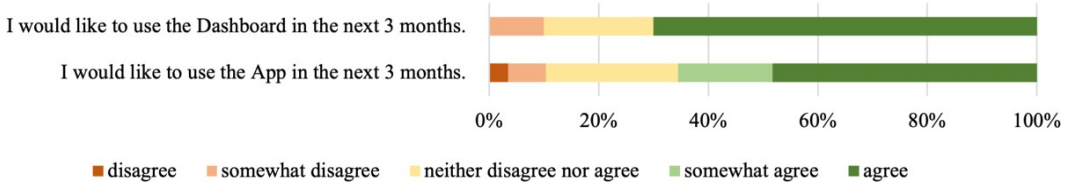


Figure 12. Behavioral Intent

Effort Expectancy: Both user groups rate the simplicity of using the system as overwhelmingly positive (Figure 13). The second item shows that the app appears to be a little easier to use than the dashboard.

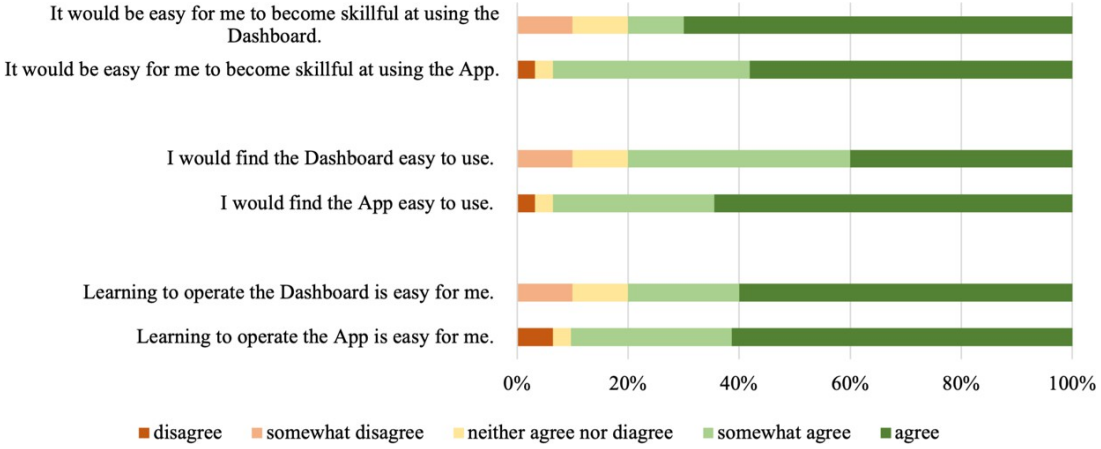


Figure 13. Effort Expectancy

Self-efficacy: Although the results for self-efficacy for both user groups are also largely positive in favor of ease of learning how to use the system, this dimension performs the weakest overall (Figure 14). It is notable that the availability of assistance would not appear to significantly improve the perception of handling the system.

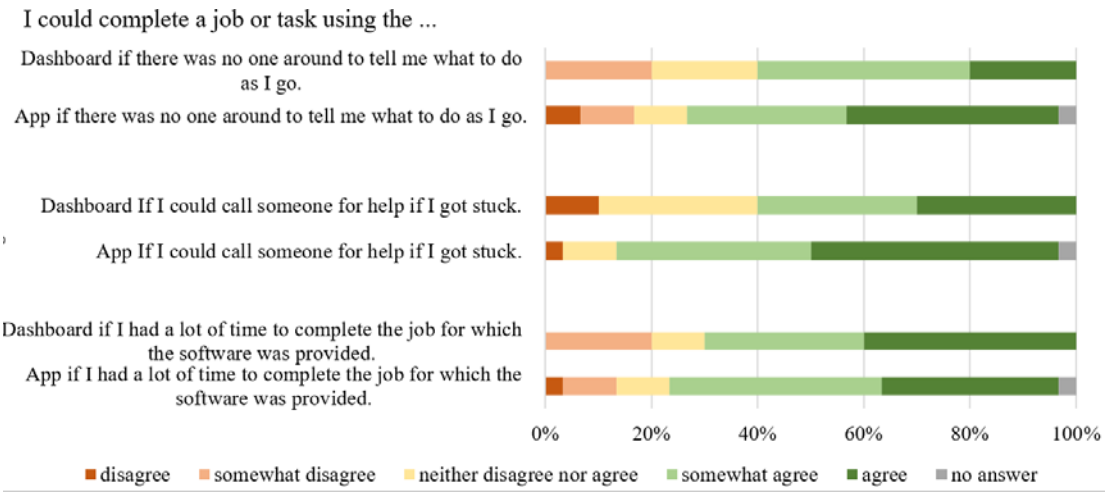


Figure 14. Self-efficacy

Satisfaction: The results indicate that both user groups are highly satisfied with the system overall (Figure 15). With a more hypothetical and less concrete formulation, the difference between app and dashboard is smaller.

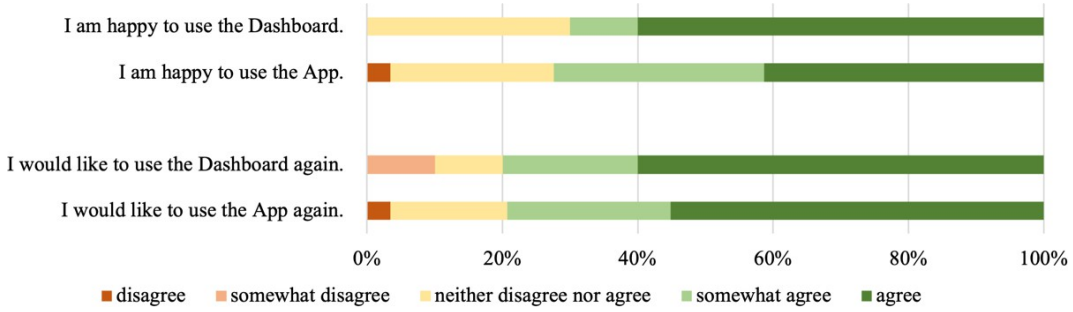


Figure 15. Satisfaction

Design: Both user groups also rated the design of their respective front ends as predominantly positive (Figure 16). The design of the app seems to be more intuitive. Still our software could be improved especially with the color codes used and in clarity and consistency of the wording.

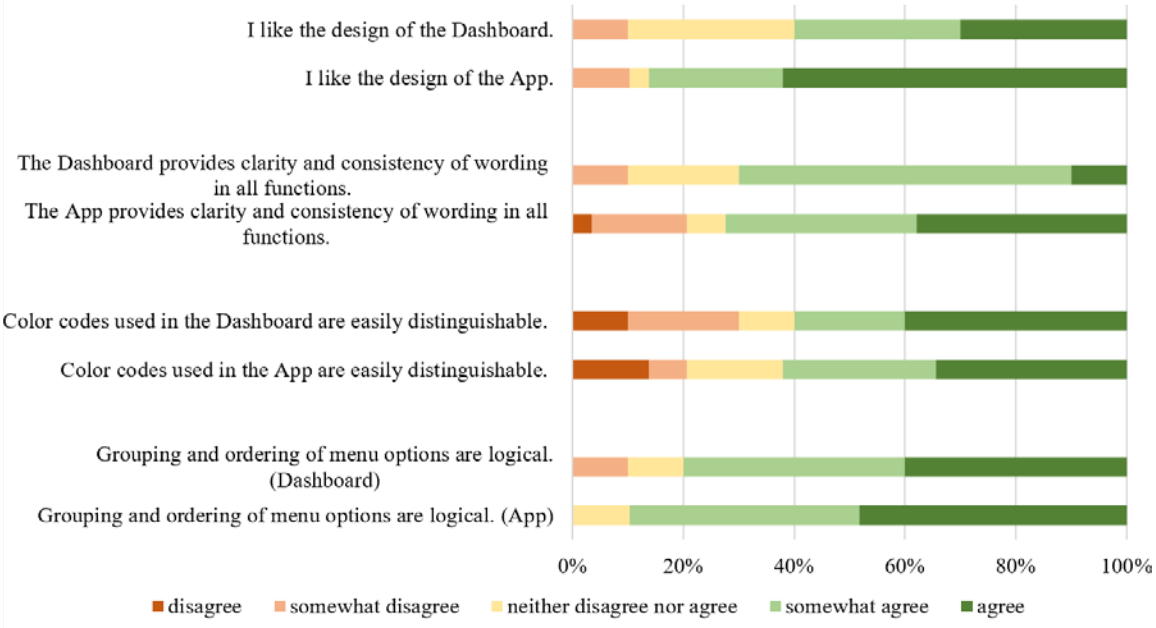


Figure 16. Design

Easiness of Use: While both user groups can easily find the information and functions, they are looking for, SVs believe that the app needs to be used regularly to be fully utilized (Figure 17). Since the answers to other dimensions, in particular effort expectancy, do not indicate that using the app is considered difficult, we assume that the users answered this question more regarding organizational aspects such as community building and gaining experience in disaster response activities and less regarding system design.

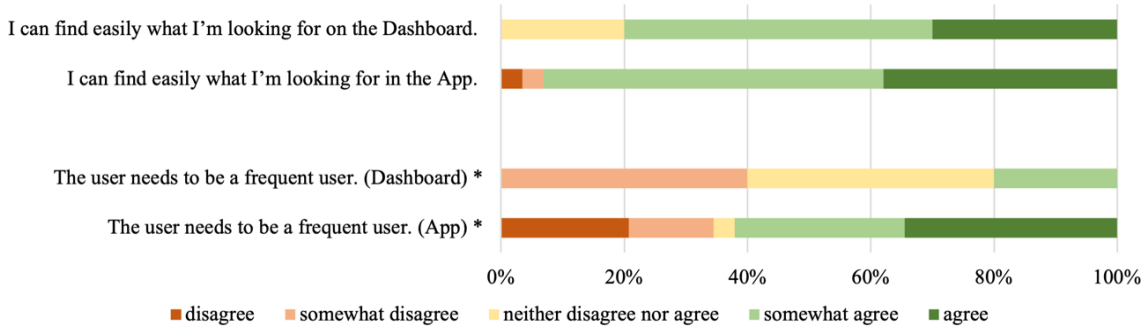


Figure 17. Easiness of Use

In summary, the results presented in relation to our second research question allow the following statements to be made: the results of the questionnaire show that most participants in both user groups have a mostly positive attitude towards the system demonstrator. Compared to the previous acceptance study (Holzhüter et al., 2023), it is clear that users appear to be much more likely to accept a real system. It is notable that the results of the SVs are more polarized and that a small group of participants is very skeptical about the system. Measured in terms of complete agreement (answers with "agree"), there is still room for improvement, particularly in terms of interface design. The good results regarding the perceived usefulness and usability of the system give us an initial indication that the theoretical results we have developed (catalog of requirements, reference architecture) can support the development of suitable systems. Due to the positive perception of the interviewees, we also assume with regard to the first research question that the development of suitable systems is possible on the basis of our design knowledge.

RELATED WORK ON FIELD TESTS OF BIDIRECTIONAL COMMUNICATION TECHNOLOGY FOR CRISES

Since we see the implementation and presentation of the evaluation of the software demonstration as the main contribution of our paper, we would like to briefly discuss the evaluation approaches of related research projects in this chapter. Although some practical studies have already been carried out in the field of volunteer coordination using social media (e.g. Batard et al.,2019; Reuter et al. 2015, Starbird 2013), the related work in relation to specific SVCS is limited. Previous field tests of bidirectional crisis communication apps either focused on an assessment of the overall importance of bidirectional communication between citizens and authorities in crises (Reuter & Kaufhold, 2018) or had a strong focus on the usability aspect of said apps (Jendreck et al., 2016).

Kaufhold et al. 2018 tested the application '112.Social', an application that allows citizens to forward alerts or report an incident to emergency services. '112.Social' then allows authorities to request additional information or disseminate behavioral tips, either via broadcasts or direct reply. Most test users evaluated the citizen to authority component positively (15 out of 21 respondents), and the same applies for the communication from authorities to citizens (14 out of 21). Furthermore, the individual app functions such as sending and receiving messages, assigning categories to messages, adding a description or a multimedia file were positively evaluated by most test persons. The only exception was the GPS tracking functionality, which was negatively evaluated, but this can be explained by the fact that the exercise took place in a building. While Kaufhold et al.'s research shows that the overall assessment of bidirectional communication in crises / disaster exercises is positive, their studies do not provide an in-depth analysis of different aspects of technology acceptance as provided by the UTAUT approach used in this paper.

The research group around the CrowdTasker system has published several papers on their approach (Auferbauer et al., 2016; Havlik et al., 2016; Middlehoff et al., 2016), each of which presents evaluation results from various experiments, although these are not mostly the focus of the paper. A significant evaluation section can be found in Auferbauer & Tellioğlu (2017), in which a field test with 12 participants, which has already been taken up in previous papers (Auferbauer et al., 2016), is discussed to a greater extent. Qualitative observations are discussed here using four categories. In Middelhoff et al. (2016), the results of a larger field test with an unknown number of participants are presented in detail. In addition to qualitative observations, a questionnaire is also used here to have users evaluate two different tools. No reference to the use of established evaluation methodology can be

found in the papers.

Jendreck et al. (2016), on the other hand, conducted a live-drill exercise with potential volunteer responders using a preliminary version of the application described in this paper. Their analysis focused on a usability assessment using the System Usability Scale (SUS). They were able to show that the application tested did have a good overall usability. Furthermore, test users positively evaluated the reliability and availability of the system tested as well as its perceived security and efficiency. The presentation of the evaluation results in (Betke et al., 2024) is similar. In this paper, the field test-based evaluation of the SVCS presented is discussed in detail, whereby the evaluation of the SVs (n = 70) was carried out quantitatively using a questionnaire and that of the disaster managers qualitatively through interviews and observations. While these studies provide important insights in technology acceptance, these studies also did not follow a holistic assessment in line with the UTAUT approach. The additional benefit provided by this paper is that it provides an overall assessment of a bidirectional crisis communication tool in a field test in line with the Unified Theory of Acceptance and Use of Technology.

CONCLUSION

In this paper we have gone through a complete design science process for the development of interconnected systems in DM. Having already conducted a requirements analysis and acceptance study and developed a reference architecture in previous work, in this paper we have focused on the implementation of a concrete software artifact in the context of the coordination of SVs and presented the demonstration and evaluation in the context of a disaster prevention exercise. We were able to show how our software demonstrator addresses the previously collected technical requirements and how its components are designed according to the reference architecture. The demonstrator was used in an exercise involving both user groups (DM, SVs) in the case of providing for large numbers of incoming refugees. The methodical evaluation based on the UTAUT approach showed that both user groups see a great benefit in the demonstrated system and perceive the implementation as very user-friendly. The results provide an initial indication that the requirements and reference architecture also provide suitable design knowledge for developing useful systems in the domain. It also indicated that a system for coordinating SVs that mees the requirements would be accepted by the majority of both user groups.

However, this is also a limitation of the approach, as the assessment based on a single use case with a small user group is only meaningful to a limited extent. However, the results are encouraging for continuing the research in further design sciences cycles. We are currently working on the implementation of an information crowdsourcing system for collecting and analyzing information from the population, which is also based on the design knowledge of our approach. Jurgens & Helsloot (2018) and Starbird (2013) have already outlined this by using social media for collective intelligence. With the evaluation of this system, we want to show that our results can be usefully applied to various interconnected systems in DM. In connection with the design of our demonstrator systems, we also collect further information to enrich our design knowledge and adapt it if necessary. The final goal of the approach is to create an overall design theory with requirements, principles, and features, which is further substantiated by an associated software architecture reference model. A final step in our approach will be to have the design knowledge evaluated by experts in the field of software development and DM to be able to make general statements about its suitability. In addition, further desiderata in the field of SVCS appear promising, so an analysis of existing approaches could synthesize the previous design knowledge into a superordinate design theory for SVCS.

ACKNOWLEDGMENTS

This work is funded by the German Federal Ministry for Economic Affairs and Climate Action under grant No. FKZ 01MK21005B SPELL.

REFERENCES

Albris, K. (2018). The switchboard mechanism: How social media connected citizens during the 2013 floods in Dresden. *Journal of Contingencies and Crisis Management, 26*(3), 350–357. https://doi.org/10.1111/1468-5973.12201

Auferbauer, D., Ganhör, R., Tellioğlu, H., & Pielorz, J. (2016). Crowdtasking: Field study on a crowdsourcing solution for practitioners in crisis management. *Proceedings of the 13th Annual Global Conference on Information Systems for Crisis Response and Management (ISCRAM 2016), May 2018*, 1–7.

Auferbauer, D., & Tellioğlu, H. (2017). Centralized crowdsourcing in disaster management: Findings and implications. *Proceedings of the 8th International Conference on Communities and Technologies*, 173–182. https://doi.org/10.1145/3083671.3083689

Batard, R., Montarnal, A., Bénaben, F., Rizza, C., Prieur, C., & Tapia, A. (2019). Integrating citizen initiatives in

a technological platform for collaborative crisis management. *Proceedings of the 16th Annual Global Conference on Information Systems for Crisis Response and Management (ISCRAM 2019)*, 1346–1356.

Betke, H., Bosse, M., Sackmann, S., & Reusch, C. (2023). Towards a Taxonomy for Classification of Coordination Systems for Spontaneous Volunteers. *Proceedings of the 20th Annual Global Conference on Information Systems for Crisis Response and Management (ISCRAM 2023)*, 751–760. https://doi.org/10.59297/bvkt8553

Betke, H., Sperling, M., Sackmann, S., & Schryen, G. (2024). A Design Theory for Spontaneous Volunteer Coordination Systems in Disaster Response. *57th Hawaii International Conference on System Science (HICSS 2024)*, 2076–2085.

Fuchs-Kittowski, F., Jendreck, M., Meissen, U., Rösler, M., Lukau, E., Pfennigschmidt, S., & Hardt, M. (2017). ENSURE - integration of volunteers in disaster management. *International Symposium on Environmental Software Systems*, 247–262. https://doi.org/10.1007/978-3-319-89935-0_21

Gerstmann, S., Betke, H., & Sackmann, S. (2019). Towards automated individual communication for coordination of spontaneous volunteers. *Proceedings of the 16th Annual Global Conference on Information Systems for Crisis Response and Management (ISCRAM 2019)*, 897–905.

Grolinger, K., Capretz, M. A. M., Mezghani, E., & Exposito, E. (2013). Knowledge as a Service Framework for Disaster Data Management. *2013 Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, 313–318. https://doi.org/10.1109/WETICE.2013.48

Havlik, D., Pielorz, J., & Widera, A. (2016). "Interaction with citizens" Experiments: From context-aware alerting to crowdtasking. *Proceedings of the 13th Annual Global Conference on Information Systems for Crisis Response and Management (ISCRAM 2016)*.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly: Management Information Systems*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Holzhüter, M., Breu, S., & Meissen, U. (2021). Use of Community Data in Crisis and Disaster Management Using the Example of Forest Fires in Germany. *EnviroInfo 2021 - Environmental Informatics – A Bogeyman or Saviour to Achieve the UN Sustainable Development Goals?*, 200–208.

Holzhüter, M., & Meissen, U. (2020). A Decentralized Reference Architecture for Interconnected Systems in Emergency Management. *Proceedings of the 17th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2020)*, 961–972.

Holzhüter, M., Reuter-Oppermann, M., Hellriegel, J., & Klafft, M. (2023). Acceptance Study on Application Systems to Improve Situational Incident Management through Bi-directional Communication between Citizens and Decision-makers in Emergencies and Crises Situations. *Proceedings of the 20th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2023)*. https://doi.org/10.59297/zlxu9879

Hwang, W., & Salvendy, G. (2010). Number of people required for usability evaluation. *Communications of the ACM*, *53*(5), 130–133. https://doi.org/10.1145/1735223.1735255

Iivari, J., & Venable, J. (2009). Action research and design science research - Seemingly similar but decisively dissimilar. *Proceeding of the 17th European Conference on Information Systems*.

Jendreck, M., Meissen, U., Fuchs-Kittowski, F., Rösler, M., & Lukau, E. (2016). ENSURE - Erhöhung der Resilienz durch die Einbindung freiwilliger Helfer in Krisensituationen. *Proceedings of the CEUR Workshop*, 129–145.

Jurgens, M., & Helsloot, I. (2018). The effect of social media on the dynamics of (self) resilience during disasters: A literature review. *Journal of Contingencies and Crisis Management*, *26*(1), 79–88. https://doi.org/10.1111/1468-5973.12212

Lin, H. X., Choong, Y. Y., & Salvendy, G. (1997). A proposed index of usability: A method for comparing the relative usability of different software systems. *Behaviour & Information Technology*, *16*(4–5), 267–277. https://doi.org/10.1080/014492997119833

Middelhoff, M., Widera, A., Van Den Berg, R. P., Hellingrath, B., Anferbauer, D., Havlik, D., & Pielorz, J. (2016). Crowdsourcing and Crowdtasking in Crisis Management Lessons Learned From a Field Experiment Simulating a Flooding in the City of the Hague. *Proceedings of the 2016 3rd International Conference on Information and Communication Technologies for Disaster Management, ICT-DM 2016*. https://doi.org/10.1109/ICT-DM.2016.7857212

Moshtari, M., & Gonçalves, P. (2017). Factors Influencing Interorganizational Collaboration within a Disaster Relief Context. *Voluntas*, *28*(4), 1673–1694. https://doi.org/10.1007/s11266-016-9767-3

Nikolai, C. M., Johnson, T., Prietula, M., Becerra-Fernandez, I., & Madey, G. R. (2015). Design Principles for Crisis Information Management Systems. *International Journal of Information Systems for Crisis Response and Management*, *7*(4), 26–45. https://doi.org/10.4018/ijiscram.2015100102

Poblet, M. (2013). Visualizing the law: crisis mapping as an open tool for legal practice. *International Workshop on AI Approaches to the Complexity of Legal Systems*, 261–274. http://ojs.law.cornell.edu/index.php/joal/article/view/12

Reuter, C. (2014). Communication between Power Blackout and Mobile Network Overload. *International Journal of Information Systems for Crisis Response and Management*, *6*(2), 38–53. https://doi.org/10.4018/ijiscram.2014040103

Reuter, C., & Kaufhold, M. A. (2018). Fifteen years of social media in emergencies: A retrospective review and future directions for crisis Informatics. *Journal of Contingencies and Crisis Management*, 41–57.

Reuter, C., Ludwig, T., Kaufhold, M. A., & Pipek, V. (2015). XHELP: Design of a cross-platform social-media application to support volunteermoderators in disasters. *Conference on Human Factors in Computing Systems - Proceedings*, 4093–4102. https://doi.org/10.1145/2702123.2702171

Reuter, C., & Spielhofer, T. (2017). Towards social resilience: A quantitative and qualitative survey on citizens' perception of social media in emergencies in Europe. *Technological Forecasting and Social Change*, *121*, 168–180.

Sakurai, M., & Murayama, Y. (2019). Information technologies and disaster management – Benefits and issues - . *Progress in Disaster Science*, *2*, 100012. https://doi.org/10.1016/j.pdisas.2019.100012

Schimak, G., Havlik, D., & Pielorz, J. (2015). Crowdsourcing in crisis and disaster management – challenges and considerations. *IFIP Advances in Information and Communication Technology*, *448*, 56–70. https://doi.org/10.1007/978-3-319-15994-2_5

Sindhuja, P. N., & Dastidar, S. (2009). Impact of the Factors influencing Website Usability on User Satisfaction. *The IUP Journal of Management Research*, *8*(12), 54–66. https://papers.ssrn.com/abstract=1524662

Starbird, K. (2013). Delivering patients to Sacré Coeur: Collective intelligence in digital volunteer communities. *Conference on Human Factors in Computing Systems - Proceedings*, 801–810. https://doi.org/10.1145/2470654.2470769

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Whitaker, J., McLennan, B. & Handmer, J. (2015). A review of informal volunteerism in emergencies and disasters: Definition, opportunities and challenges. *International Journal of Disaster Risk Reduction*, 13, 358-368.

The coordination and leadership in first aid ad-hoc immediate responder groups: a work-in-progress

Wilhelm Brodin
Department of Computer and Information Science, Linköping University
wilhelm.brodin@liu.se

Carl-Oscar Jonson
Center for Disaster Medicine and Traumatology, Linköping University
carl-oscar.jonson@regionostergotland.se

Erik Prytz
Department of Computer and Information Science, Linköping University
erik.prytz@liu.se

ABSTRACT

When an accident has occurred, first aid before a professional response arrives can improve the outcome for the victims. However, little research has been done on the effect of coordination and leadership on first aid responses. This work-in-progress study reports an experiment using a simulated scenario, i.e. traffic accident, to understand the effect of coordination and leadership on first aid performance in ad-hoc groups of immediate responders. The results showed that coordination and leadership are significantly related to increased self-evaluated team performance while expertise and authority differentiation is not. These results provide a first insight to the coordination and leadership in ad-hoc groups of immediate responders, but more research is needed to validate these results. Structured performance evaluation with subject matter experts is planned as an extension to the present work-in-progress.

Keywords

First aid, immediate responders, emergent groups, leadership, coordination.

INTRODUCTION

Two major health emergencies in the world are out-of-hospital cardiac arrests (OHCA) and traumatic injuries. Over 10 000 people in Europe (Gräsner et al., 2016) and over 30 000 in the US (Virani et al., 2020) suffer from OHCA each month, and traumatic injuries are one of the leading causes of death for young adults globally (Vos et al., 2020). Immediate responders often provide first aid to people who suffer from these medical emergencies, starting cardiopulmonary resuscitation (CPR) in 58% of CPR cases (Gräsner et al., 2016) and providing first aid for traumatic injuries in 59% of traumatic injury cases (Pelinka et al., 2004). An immediate responder is a person that happens to be present where someone needs first aid and provides aid before the professional responders arrive (Harris et al., 2018). Immediate responder intervention increases the chance of survival for the injured person in both OHCA and traumatic injuries such as massive bleeding (Gräsner et al., 2016; Hashmi et al., 2022). Often there are multiple immediate responders providing aid simultaneously (Bakke et al., 2015), which could increase the quality of the aid provided (Takei et al., 2014).

Accidents and OHCA can happen wherever people are, and which individuals will be close and provide aid is therefore random. The immediate responders who decides to provide aid may not know each other beforehand and they would have to quickly establish trust (i.e. swift or immediate trust; Olsen et al., 2020) for an effective response. Furthermore, the first aid knowledge among the immediate responders might also vary. Bakke et al.

(2015) found that many of the immediate responders to first aid traumatic injuries were individuals with jobs such as healthcare professional, police officer and firefighter, but were responding outside of their professional role. While it is likely that they have knowledge about correct first aid measures, if the immediate responders do not know each other beforehand it can be difficult to quickly learn each individual's level of knowledge in first aid. Even in a group with higher levels of knowledge, odds are that they likely have never trained in providing first aid together, which could negatively affect the group's performance (Burke et al., 2004).

Research on ad-hoc groups of immediate responders providing first aid is scarce but has been studied in disasters through case studies (Stallings & Quarantelli, 1985; Twigg & Mosel, 2017; Drabek & McEntire, 2003). Although the scope of the activities of immediate responder in ad-hoc groups in disasters is larger than that of the typical first aid situation, there are similarities between them that make the literature useful for initial expectations of the coordination and leadership. For example, the task of providing first aid by groups of varying and unknown levels of knowledge is present in both disasters and everyday accidents. In disasters, the organization and coordination of ad-hoc groups of immediate responders is described as having a flat hierarchy with loose membership and driven by a mutual feeling of togetherness in the group (Stallings & Quarantelli, 1985; Twigg & Mosel, 2017). The flat hierarchy means that there is no particular individual that is given or takes the leadership role. This flat hierarchy is also known as low authority differentiation (Hollenbeck et al., 2012). However, that does not necessarily mean that the groups have no leadership within them. Leadership could instead be conceptualized as a function (i.e. functional leadership) that is distributed between the individuals in the group rather than a role assigned to a specific individual in the group (Morgeson et al. 2010). The loose membership of an ad-hoc group means that the membership of the group is not predetermined and is subject to change during a response.

Majchrzak et al. (2007) provides a suggestion to extend the theory of transactive memory systems (TMS) so that ad-hoc groups of immediate responders in disasters can be understood as a TMS. A group that works together interdependently toward a common goal is defined as a TMS if the individuals in the group have differentiated knowledge and uses a dynamically coordinated process within the group to encode, store and retrieve knowledge (Lewis & Herndon, 2011). Coordination in the context of TMS can thus be conceptualized as the smooth adaption of actions by individuals to interdependently achieve a common goal (Lewis, 2003). The TMS structure and processes in a group improves team performance (DeChurch & Mesmer-Magnus, 2010) but there are moderating factors, such as leadership, that could influence the effect (Bachrach et al., 2019). Majchrzak et al. (2007) further suggested that expertise in ad-hoc immediate responder groups is coordinated through actions, such that experts communicate their higher levels of knowledge and ability by initiating a response and the non-experts are then to understand the response as a sign of that individual's greater knowledge-level. Additionally, Majchrzak and colleagues suggest that individuals with less knowledge learn through action within the response, i.e. learning by doing.

The lack of experimental research on ad-hoc groups of immediate responders and the implied particularity of such groups made by Majchrzak et al. (2007) suggests a need for further understanding of these groups in more controlled circumstances. Therefore, this work in progress-study aims to use experimental simulation to investigate how aspects of coordination, and leadership affect the performance of ad hoc groups of immediate responders.

METHODS

An experiment with a simulated traffic accident was conducted with a between-group design. Groups of three participants participated, and groups were classified as either low knowledge or high knowledge. The low knowledge groups included three medical laypeople, i.e. individuals with little or no medical knowledge or experience of first aid. In contrast, the high knowledge groups included two medical laypeople and one first responder expert. The experts had previous education and experience in pre-hospital healthcare or as an emergency responder. A total of 28 groups participated in the experiment. However, 9 groups were excluded from the analysis due to participant no-shows resulting in group sizes below three. See Table 1 for demographics of the participants included in the analysis.

To preserve the ad-hoc aspect of the simulation, the participants were not informed that multiple participants would participate in the same scenario. The study was therefore considered a deception study by the Swedish Ethical Review Authority and ethical approval was acquired. Participants received monetary compensation for participation.

Table 1. Demographics of the participants included in the analysis.

		Low knowledge teams			High knowledge teams			Full sample		
		n	%		*n*	%		*n*	%	
Teams (groups of three participants)		12	63		7	37		19	100	
Participants (individuals)		36	63		21	37		57	100	
Gender										
	Female	15	42		6	29		21	37	
	Male	20	55		14	66		34	60	
	Prefer to not answer	1	3		1	5		2	3	
Level of education										
	Compulsory school	0	0		1	5		1	2	
	High school diploma	8	22		3	14		11	19	
	A few years at university	16	45		10	48		26	46	
	Bachelor's degree	7	19		4	19		11	19	
	Master's degree	5	14		3	14		8	14	
	PhD	0	0		0	0		0	0	
First aid education										
	Yes	18	50		16	76		34	60	
	No	18	50		5	24		23	40	
		n	M	SD	*n*	M	SD	*n*	M	SD
Years since first aid education		15	5.1	3.6	15	5.6	9.0	30	5.3	6.7
Age		36	24.7	7.2	21	31.5	13.6	57	27.2	10.5

Note. High knowledge teams included one first responder expert. Level of education is the participants highest achieved level of education at the time of their participation. 4 participants with previous first aid education did not disclose years since they received the education.

Procedure

A convenience sample of medical laypeople was recruited through posters at a university in southern Sweden. The emergency response (prehospital or rescue service) experts were recruited by providing information about the study to relevant employers in the immediate vicinity of the experiment location (see Figure 1 for procedure flowchart). Potential participants that registered for participation were vetted according to medical education and first aid experience. Participants with emergency response expertise were specifically asked to avoid wearing clothes showing text or logos indicating their professional roles to avoid implicitly communicating their expertise to the other participants. Participants who might know each other beforehand (e.g., studied in the same program) were scheduled to different time slots to preserve the ad-hoc aspect of the groups.

At the start of the experiment, the participants were brought one by one to a room where they were separately seated and were asked to read and sign an informed consent form. Sightlines between the seats were blocked, and the participants also wore hearing protection so that they could not hear instructions or questions from other participants. Participants were fitted with a body cam and a lavalier microphone. The participants were individually informed that they could stop the scenario at any point and instructed to act as they thought they would if the scenario was real. To disambiguate their role in the scenario the instructions included contextual information about the scenario. Specifically, that they had heard a car crash and that they had decided to investigate and provide aid in any way they were able to. This was motivated by the aim of the experiment to study first aid response rather than the decision to initiate first aid.

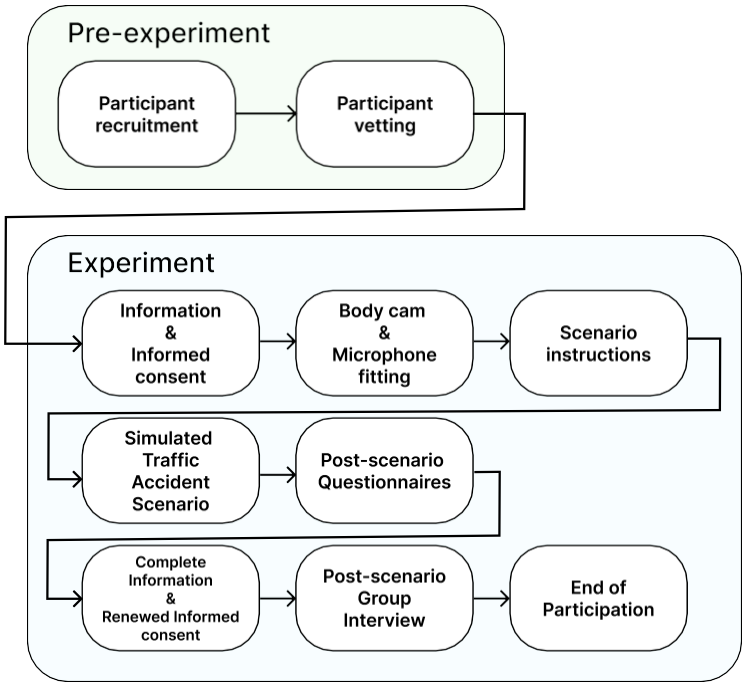


Figure 1. Procedure flowchart

The participants started the scenario simultaneously and the scenario took a maximum of 15min. It included a car with gray smoke from the engine bay and two injured actors sitting in the car. The actor closest to the participants was a human amateur actor with a simulated massive bleeding from the left upper arm. All actors received 2h of training to ensure that they could portray the injury in similar ways with regards to consciousness and responsiveness. The actors were instructed to show drowsiness, to be slow to respond to light stimulation similar to level 2 or 3 on the Reaction Level Scale (Starmark et al., 1988), and to be unable to walk. A protective sleeve with a silicone wound connected to a 1.5L bottle of fake blood was used to simulate the bleeding in the actors' left arm (Tourniquet application trainer, 2023). The flow of the blood was controlled by the actors, who pumped blood until bleeding control via first aid had been achieved. An additional 500ml of fake blood was poured close to the initial position of the actor before the start of the scenario to indicate that the bleeding was severe. The other injured actor consisted of a rescue manikin (Rescue manikin Pro; Svenska Räddningsprodukter AB., n.d) modified to enable CPR compressions (see Figure 2). The rescue manikin had a realistic weight distribution and weighed 80kg, realistic face and arm features in silicone, and no detectable breathing or pulse. Participants that were visibly attempting to detect the manikin's pulse or breathing were prompted by the research leader that they did not detect any pulse or breathing. Groups that decided to simulate a call to the Swedish emergency dispatch center (i.e. SOS Alarm) got a response from the research leader simulating the dispatch center response. After a participant provided the information they deemed important the research leader ended the simulated call by informing the participant that professional help was on its way.

A small red first aid pouch (First Aid pouch, mini, n.d.) was included in the scenario. It included a rescue blanket, non-hemostatic gauze, one pair of plastic gloves, a face shield for rescue breaths and disinfectant wipes. Other minor items such as band aids, scissor and a gel packet for burns were also included in the pouch. A combat application tourniquet model CAT7 was placed in the pouch to enable participants to effectively stop a massive bleed in extremities. In Sweden, there is no legal obligation to keep first aid material in the car and participants may not expect to find such material in a car. Therefore, the first aid pouch was placed in the open glove compartment of the car, visible when the passenger door was opened

After the scenario the participants were seated separately and completed post-scenario questionnaires (see Materials section for a description). After the questionnaires, the participants were seated together and provided complete information about the purpose of the study and informed consent was renewed. Then a group interview focusing on the participants' experience of leadership and coordination during the scenario was conducted. The interview took approximately 30-50min and was audio recorded. The participants had the choice to conduct the interview separately as they had not received information about the interview being conducted in a group beforehand. Lastly, the experiment was concluded by thanking the participants for their participation.



Figure 2. Simulated traffic accident scenario

Materials

Data from the following questionnaires are included in the current work-in-progress analysis. A demographics questionnaire collected data on the age, gender, level of education, first aid education, number of previous first aid experiences, current or previous work as an emergency responder, and previous education and experience of teamwork and leadership. Coordination was measured through a questionnaire on the coordination dimension of TMS (Lewis, 2003). Leadership was measured using an adaption of the questionnaire used in a previous study on teamwork in a simulated pre-hospital scenario (Hughes et al., 2021) and an adaption of the questionnaires on functional leadership in transition processes (Morgeson et al., 2010). Transition processes in a team are the actions that team members do during an entire collaboration to make a plan and mission analysis, specify goals for the team and formulate team strategies (Marks et al., 2001). The participants' experienced authority differentiation was measured by adapting the description of authority differentiation by Hollenbeck et al. (2012) into two questions about leadership experienced as centralized or distributed. Data on participants' self-evaluated performance regarding team performance, leadership performance and individual performance was also collected. Note that only the self-evaluated team performance data was used in the current work-in-progress analysis.

The following questionnaire data was collected but not analyzed further in the current work-in-progress. The participants' experienced team processes were collected using a questionnaire based on the temporal taxonomy of team processes (Mathieu et al., 2020; Marks et al., 2001). A questionnaire on team orientedness was collected as a measure for participants' general tendency to prefer working in groups over working alone (Driskell et al., 2010). Two questions regarding participants' general and immediate trust were adapted from Delhey et al. (2011) and Olsen et al. (2020).

Video was recorded using a camera with 360° field of view and five small cameras: one strapped to each participant, one placed inside of the cabin of the car and one operated by the research leader. The audio of the participants' cameras was recorded using lapel microphones. Audio recordings were made during the group interviews using a field recorder. The group interview was a semi-structured interview with questions based on the questionnaires about functional leadership (Morgeson et al., 2010), team processes (Mathieu et al., 2020) and coordination (Lewis, 2003). During the group interviews the participants were asked to share their experiences related to leadership and coordination during the scenario.

Analysis methods

A two-level hierarchical model was used to assess the effects of expertise, coordination, authority differentiation and functional leadership in transition processes on self-assessed team performance (see Table 2 for a description of the variables). The model was constructed in R Statistical software (v 4.2.1; R Core Team, 2022) using the nlme package (v 3.1.157; Pinheiro et al., 2022). All predictors were on an individual level and thus included in the model as first level units and the team that the participants acted in the scenario with was used as the second-level unit. Only teams with three participants were included in the analysis, resulting in 57 participants in 19 teams. The multilevel model method was used to avoid the assumption of independence in linear regressions.

Distributions for predictors were acceptable and no outliers were found. All predictors were modelled with random intercepts and fixed slopes due to limited amount of data resulting in models with random slopes not converging. The random intercept and fixed slopes in the model mean that the relation between the predictors and the outcome is the same for all teams but the intercept can vary between teams. As the current work is a work-in-progress the analysis did not include any statistical analysis concerning the final model's generalizability.

Table 2. Description of variables included in the analysis.

Concept	Operationalization	Measurement	Reference
Expertise	Current or previous work as an emergency responder	Binary coding of previous/current experience or no previous/current experience as emergency responder	
Coordination	Coordination in TMS	Mean score of 5 questions on a Likert-scale (1-5) with the third question reverse coded	(Lewis, 2003)
Leadership	Functional leadership in transition processes	Mean score of 7 questions on a Likert-scale (1-5)	(Hughes et al., 2021; Morgeson et al., 2010)
	Authority differentiation	Mean score of 2 questions on a Likert-scale (1-5) with the first question reverse coded	(Hollenbeck et al., 2012)
Performance	Team performance	1 question of self-evaluation of the team performance on a Likert-scale (1-5)	

PRELIMINARY RESULTS

The two-level hierarchical model showed coordination and functional leadership in transition processes as significant predictors for self-assessed team performance, while expertise and authority differentiation did not reach significance as predictors (see Table 3). Expertise was found to have a positive effect on self-assessment of team performance. Meaning that an expert participant in this sample tended to rate the team performance higher than a non-expert participant. However, the effect was not statistically significant ($p = 0.053$) with the current sample size ($n = 57$). Therefore, further data collection is needed to determine if the effect is generalizable. Coordination significantly predicted self-assessment of team performance which indicates that participants who experienced more coordination within the group tended to rate the team performance higher. Authority differentiation had a negative effect on self-assessed team performance which means that participants in the current sample that experienced more equal authority distribution in the group tended to rate the team performance higher. However, the effect was not statistically significant ($p = 0.051$) with the current sample size ($n = 57$) and further data collection is needed to determine the generalizability of the authority differentiation effect as well. Functional leadership in transition processes was found to significantly predict self-assessed team performance. Participants that experienced more leadership functions related to team processes that involve planning and understanding the available resources within the group tended to rate team performance higher.

Table 3. The fixed effects of expertise, coordination, authority differentiation and functional leadership in transition processes on self-assessed team performance

	b	SE *b*	DF	t-statistic	*p*	95% CI
Expertise	0.51	0.25	34	2.00	.053	0.02, 1.00
Coordination	0.73	0.20	34	3.59	.001	0.34, 1.13
Authority differentiation	-0.15	0.08	34	-2.02	.051	-0.30, -0.01
Functional leadership in transition processes	0.48	0.13	34	3.66	<.001	0.22, 0.73

DISCUSSION

The result from the current study shows that coordination is positively related to self-evaluated team performance in ad-hoc first aid groups. Coordination has in previous research also been shown to be important for performance in other types of teams (Rico et al., 2019). It would therefore be reasonable to assume that coordination is important for team performance in ad-hoc groups of immediate responders. Additionally, the tasks in the current

scenario were designed to be too difficult to complete successfully for one single person. For example, continuing CPR while also stopping a bleed is impossible for one person as they have a limited number of hands and a limit to their capacity to multitask. However, as the performance measure in the current study consists of self-evaluation rather than a more objective or structured measurement it is also possible to interpret the positive effect of coordination on team performance as coordination being part of the participants mental model of what constitutes more effective team performance. Further analysis is warranted to investigate the effect of coordination on the performance of the first aid interventions attempted by the participants rather than their ability to work together.

Functional leadership in transition processes was also found to positively contribute to the self-assessed team performance in the simulated scenarios. In the current study leadership was defined as a function that one or more people could fill rather than a role that one individual was assigned. This was done due to the emergent and ad-hoc nature of organization of immediate responder groups (Drabek & McEntire, 2003). Dagnell (2020) suggests some characteristics of leadership that have been found to positively influence performance in intra-hospital resuscitation simulations that could also influence the quality of CPR in OHCA. Those overlapping with the operationalization of leadership in the current study are task assignment, planned tasks, and structure initiation. The result from the current study supports part of the suggestion by Dagnell (2020) in the context of CPR by immediate responders in an OHCA simulation. However, (Dagnell, 2020) concerns ad-hoc teams of medical professionals in a pre-hospital context which could have more resources and knowledge available compared to the ad-hoc groups of immediate responders in the current study.

Pelinka et al. (2004) found that immediate responders with more medical education tend to give better aid to victims with traumatic injuries. In the current study, expertise showed a similar effect but the effect did not reach statistical significance. Based on the results reported by Pelinka et al. (2004), expertise would be expected to predict self-evaluated team performance as well. Quarantelli (1988) reports that inability to change strategy and structure in the face of new and unprecedented change negatively affects performance in disaster response. Additionally, difference in team cognition could also be negatively affecting performance (Mohammed et al., 2021). An insufficient overlap in relevant content (knowledge) and procedures (strategy) between experts and medical laypeople could create conflicts that take time and resources from achieving the task at hand. However, the use of self-evaluation of performance as a proxy for a performance measurement could also explain the results as a mismatch between what medical laypeople perceives as good performance and what constitutes actual performance in a first aid setting.

Authority differentiation was not a significant predictor for self-evaluated performance in the current study. In previous research, similarity in goal orientation within a team has been shown to moderate the positive effect of authority differentiation on team performance (Nederveen Pieterse et al., 2019). Specifically, a team where individuals' goal prioritization differs more would increase performance in teams with greater authority differentiation. Interpreting the negative effect that authority differentiation tended toward with regards to team performance in the current study through the interaction described by Nederveeen Pieterse et al. (2019) suggests that there was disagreement within the teams about which goals to pursue. However, the content of the participants goals was not collected in the current study making direct comparison between the results in the current study and the study by Nederveeen Pieterse et al. (2019) difficult.

Limitations

The current study is one of the first to investigate coordination and leadership in ad-hoc immediate responder teams in a simulated scenario. The simulated scenario design allows for more control and reliability compared to the case study methodology used in previous work on ad-hoc immediate responder groups in disasters and crisis management. The experimental simulation methodology also provides more ecological validity than laboratory experiments (McGrath, 1981). However, the study contains some noteworthy limitations. The current study is a work-in-progress and is limited by the use of self-assessment of team performance as outcome variable. Previous research indicates that self-evaluation of performance is only moderately correlated to actual performance (Zell & Krizan, 2014). Therefore, a structured performance evaluation with subject matter experts is being undertaken to gain a more accurate measure of the teams' performance in the simulated scenario (see Future work section).

Although over 70 individuals participated in the study the number of teams included in the analysis was limited. Groups of three participants were planned but due to participant no-shows a third of the scenarios included two participants and were thus excluded from the current analysis. Participant no-shows resulted in fewer datapoints than expected and therefore decreases the chance of statistically determining effects of predictors on the outcome variable. The expertise and authority differentiation variables may suffer from such decreased effects in the current study and further data collection is warranted. However, the results provide a valuable new methodological perspective to a problem previously mainly explored through case studies.

Future work – Performance evaluation procedure

Future work currently under way aims to mitigate the self-evaluated team performance limitation by employing a structured evaluation process using performance indicators and subject matter experts. Performance indicators have been selected in a workshop with two subject matter experts with expertise in pre-hospital healthcare and rescue and response expertise. During the workshop, video and audio from the pilot was used to exemplify which demands the scenario placed on the participants to match the performance indicators with the scenario. Five performance indicators were selected: first aid performance for CPR and hemorrhage control, contacting emergency dispatch, maintaining safety on the scene of the accident and communication with the conscious injured person. The indicators are also designed to measure task performance and not coordination or leadership. Each performance indicator is rated on a scale 1-7 by subject matter experts using the video and audio data collected during the trials. The video and audio data have been edited into a single video to make the rating process easier. After rating each video an interrater measurement will be used to ensure unbiased performance ratings. The swift trust data and the group interview recordings are planned to be included in future work by comparing how participants describe their experience of coordination and leadership with the proposed extension of TMS by Majchrzak et al. (2007).

CONCLUSION

Coordination and functional leadership in transition processes was shown to predict self-evaluated team performance. Expertise and authority differentiation was also found to influence self-evaluated team performance, but the effects did not reach statistical significance. This suggests that more is needed for effective first aid in ad-hoc groups of immediate responders than making sure that there is first responder experience and one individual acting as a leader. Coordination and functional leadership as significant predictors also suggest that they are a vital part of successful first aid. As such, more work is needed to understand how these factors could be included into first aid education to improve immediate responders' ability to provide effective aid in the future. This work-in-progress study also contributes to the understanding of how successful first aid is conducted in a difficult first aid scenario by using a research method with more control and reliability compared to previous research on immediate responders. Future work aims to investigate the effect of coordination and leadership on performance as rated by subject matter experts.

REFERENCES

Bachrach, D. G., Lewis, K., Kim, Y., Patel, P. C., Campion, M. C., & Thatcher, S. M. B. (2019). Transactive memory systems in context: A meta-analytic examination of contextual factors in transactive memory systems development and team performance. Journal of Applied Psychology, 104(3), 464–493. https://doi.org/10.1037/apl0000329

Bakke, H. K., Steinvik, T., Eidissen, S. -I., Gilbert, M., & Wisborg, T. (2015). Bystander first aid in trauma – prevalence and quality: A prospective observational study. Acta Anaesthesiologica Scandinavica, 59(9), 1187–1193. https://doi.org/10.1111/aas.12561

Burke, C., Salas, E., Wilson-Donnelly, K., & Priest, H. (2004). How to turn a team of experts into an expert medical team: Guidance from the aviation and military communities. Quality & safety in health care, 13(Suppl 1), 96–104. https://doi.org/10.1136/qshc.2004.009829

Dagnell, A. J. (2020). Teamwork and Leadership in Out-of-Hospital Cardiac Arrest – Do These Non-Technical Skills Require Attention? Australasian Journal of Paramedicine, 17, 1–6. https://doi.org/10.33151/ajp.17.748

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology*, *95*(1), 32–53. https://doi.org/10.1037/a0017328

Delhey, J., Newton, K., & Welzel, C. (2011). How General Is Trust in "Most People"? Solving the Radius of Trust Problem. American Sociological Review, 76(5), 786–807. https://doi.org/10.1177/0003122411420817

Drabek, T. E., & McEntire, D. A. (2003). Emergent phenomena and the sociology of disaster: Lessons, trends and opportunities from the research literature. Disaster Prevention and Management: An International Journal, 12(2), 97–112. https://doi.org/10.1108/09653560310474214

Driskell, J. E., Salas, E., & Hughes, S. (2010). Collective Orientation and Team Performance: Development of an Individual Differences Measure. Human Factors: The Journal of the Human Factors and Ergonomics Society, 52(2), 316–328. https://doi.org/10.1177/0018720809359522

First Aid pouch, mini. (n.d.). Biltema.se. https://www.biltema.se/en-se/car---mc/mc/first-aid-kits/first-aid-pouch-mini-2000031346

Gräsner, J.-T., Lefering, R., Koster, R. W., Masterson, S., Böttiger, B. W., Herlitz, J., Wnent, J., Tjelmeland, I. B. M., Ortiz, F. R., Maurer, H., Baubin, M., Mols, P., HadžibegoviĿ, I., Ioannides, M., Škulec, R., Wissenberg, M., Salo, A., Hubert, H., Nikolaou, N. I., … Whittington, A. (2016). EuReCa ONE - 27 Nations, ONE Europe, ONE Registry: A prospective one month analysis of out-of-hospital cardiac arrest outcomes in 27 countries in Europe. *Resuscitation*, *105*, 188–195. https://doi.org/10.1016/j.resuscitation.2016.06.004

Harris, C., McCarthy, K., Liu, E., Klein, K., Swienton, R., Prins, P., & Waltz, T. (2018). Expanding Understanding of Response Roles: An Examination of Immediate and First Responders in the United States. International Journal of Environmental Research and Public Health, 15(3), 534. https://doi.org/10.3390/ijerph15030534

Hashmi, Z. G., Hu, P. J., Jansen, J. O., Butler, F. K., Kerby, J. D., & Holcomb, J. B. (2022). Characteristics and Outcomes of Prehospital Tourniquet Use for Trauma in the United States. Prehospital Emergency Care, 1–7. https://doi.org/10.1080/10903127.2021.2025283

Hollenbeck, J. R., Beersma, B., & Schouten, M. E. (2012). Beyond Team Types and Taxonomies: A Dimensional Scaling Conceptualization for Team Description. Academy of Management Review, 37(1), 82–106. https://doi.org/10.5465/amr.2010.0181

Hughes, A. M., Sonesh, S. C., Mason, R. E., Gregory, M. E., Marttos, A., Schulman, C. I., & Salas, E. (2021). Trauma, Teams, and Telemedicine: Evaluating Telemedicine and Teamwork in a Mass Casualty Simulation. Military Medicine, 186(7–8), e811–e818. https://doi.org/10.1093/milmed/usaa434

Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. Journal of Applied Psychology, 88(4), 587–604. https://doi.org/10.1037/0021-9010.88.4.587.

Lewis, K., & Herndon, B. (2011). Transactive Memory Systems: Current Issues and Future Research Directions. *Organization Science*, *22*(5), 1254–1265. https://doi.org/10.1287/orsc.1110.0647

Majchrzak, A., Jarvenpaa, S. L., & Hollingshead, A. B. (2007). Coordinating Expertise Among Emergent Groups Responding to Disasters. Organization Science, 18(1), 147–161. https://doi.org/10.1287/orsc.1060.0228

Marks, M. A., Mathieu, John. E., & Zaccaro, S. J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. The Academy of Management Review, 26(3), 356–376. https://doi.org/10.2307/259182

Mathieu, J. E., Luciano, M. M., D'Innocenzo, L., Klock, E. A., & LePine, J. A. (2020). The Development and Construct Validity of a Team Processes Survey Measure. Organizational Research Methods, 23(3), 399–431. https://doi.org/10.1177/1094428119840801

McGrath, J. E. (1981). Dilemmatics: The Study of Research Choices and Dilemmas. American Behavioral Scientist, 25(2), 179-210. https://doi.org/10.1177/000276428102500205

Mohammed, S., Rico, R., & Alipour, K. K. (2021). Team Cognition at a Crossroad: Toward Conceptual Integration and Network Configurations. Academy of Management Annals, 15(2), 455–501. https://doi.org/10.5465/annals.2018.0159

Morgeson, F. P., DeRue, D. S., & Karam, E. P. (2010). Leadership in Teams: A Functional Approach to Understanding Leadership Structures and Processes. Journal of Management, 36(1), 5–39. https://doi.org/10.1177/0149206309347376

Nederveen Pieterse, A., Hollenbeck, J. R., van Knippenberg, D., Spitzmüller, M., Dimotakis, N., Karam, E. P., & Sleesman, D. J. (2019). Hierarchical leadership versus self-management in teams: Goal orientation diversity as moderator of their relative effectiveness. The Leadership Quarterly, 30(6), 101343. https://doi.org/10.1016/j.leaqua.2019.101343

Olsen, O. K., Heesch, P. v., Søreide, C., & Hystad, S. W. (2020). Trust After Just 45 Seconds? An Experimental Vignette Study of How Leader Behavior and Emotional States Influence Immediate Trust in Strangers. Frontiers in Psychology, 10, 2921. https://doi.org/10.3389/fpsyg.2019.02921

Pelinka, L. E., Thierbach, A. R., Reuter, S., & Mauritz, W. (2004). Bystander trauma care—Effect of the level of training. Resuscitation, 61(3), 289–296. https://doi.org/10.1016/j.resuscitation.2004.01.012

Pinheiro J., Bates D., & R Core Team. (2022). _nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-157. https://CRAN.R-project.org/package=nlme

Quarantelli, E. L. (1988). Disaster Crisis Management: A Summary of Research Findings. Journal of Management Studies, 25(4), 373–385. https://doi.org/10.1111/j.1467-6486.1988.tb00043.x

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rico, R., Gibson, C. B., Sánchez-Manzanares, M., & Clark, M. A. (2019). Building team effectiveness through adaptation: Team knowledge and implicit and explicit coordination. Organizational Psychology Review, 9(2–3), 71–98. https://doi.org/10.1177/2041386619869972

Stallings, R. A., & Quarantelli, E. L. (1985). Emergent citizen groups and emergency management. Public administration review, 45, 93-100.

Starmark, J. E., Stålhammar, D., & Holmgren, E. (1988). The reaction level scale (RLS 85) manual and guidelines. Acta neurochirurgica, 91, 12-20.

Svenska Räddningsprodukter AB. (n.d.). Rescue Manikin Pro: Svenska räddningsprodukter AB. Svenska räddningsprodukter AB, Sala, Sweden. https://en.svenskaraddningsprodukter.se/products/rescue-manikin-pro

Takei, Y., Nishi, T., Matsubara, H., Hashimoto, M., & Inaba, H. (2014). Factors associated with quality of bystander CPR: The presence of multiple rescuers and bystander-initiated CPR without instruction. Resuscitation, 85(4), 492–498. https://doi.org/10.1016/j.resuscitation.2013.12.019

Tourniquet Application Trainer. (2023). TrueClot. https://trueclot.com/product/tourniquet-application-trainer/

Twigg, J., & Mosel, I. (2017). Emergent groups and spontaneous volunteers in urban disaster response. Environment and Urbanization, 29(2), 443–458. https://doi.org/10.1177/0956247817721413

Virani, S. S., Alonso, A., Benjamin , E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Chang, A. R., Cheng, S., Delling, F. N., Djousse, L., Elkind, M. S. V., Ferguson, J. F., Fornage, M., Khan, S. S., Kissela, B. M., Knutson, K. L., Kwan, T. W., Lackland, D. T., … on behalf of the American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2020) Heart disease and stroke statistics—2020 update: a report from the American Heart Association. Circulation, 141, e139–e596. https://doi.org/10.1161/CIR.0000000000000757Circulation.

Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., Abdollahi, M., Abdollahpour, I., Abolhassani, H., Aboyans, V., Abrams, E. M., Abreu, L. G., Abrigo, M. R. M., Abu-Raddad, L. J., Abushouk, A. I., … Murray, C. J. L. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. The Lancet, 396(10258), 1204–1222.

https://doi.org/10.1016/S0140-6736(20)30925-9

Zell, E., & Krizan, Z. (2014). Do People Have Insight Into Their Abilities? A Metasynthesis. Perspectives on Psychological Science, 9(2), 111–125. https://doi.org/10.1177/1745691613518075

Selection of rural villages as volunteer centers for emergency response

Niki Matinrad
Linköping University,
Department of Science and Technology
niki.matinrad@liu.se

Tobias Andersson Granberg
Linköping University,
Department of Science and Technology
tobias.andersson.granberg@liu.se

ABSTRACT

Rural areas are typically sparsely populated, and coverage from emergency services is weak. In some such areas in Sweden, professional emergency services train civilian volunteers in providing help for certain types of emergencies. In case of an emergency, they are alerted on their mobile phones and are primarily expected to provide help in the vicinity of their homes. Due to limited resources, it is impossible to include all villages in the volunteer initiative, making the selection of the most beneficial villages for emergency response important. However, this decision is not always trivial to make. Therefore, we propose an optimization model to determine which villages should be included in this initiative to achieve the best possible emergency response coverage. The model is used in a case study for a large Swedish fire and rescue service organization, and based on the results, they have started volunteer centers in two villages.

Keywords

Volunteers, Civilian first responders, Optimization, Coverage, Emergency response.

INTRODUCTION

Emergency services have a pivotal role for managing emergencies ranging from daily emergencies (frequent events that have low consequences such as traffic accidents (Quarantelli, 1995)) to disasters, which happen seldomly but have high level of consequence. To manage any type of emergency, these organizations use different resources, including personnel, physical equipment, and information. However, the availability of these resources plays an important role in their service delivery (Matinrad, 2022). In recent years, emergency services have been facing different challenges, including cost pressures and budget cutbacks as well as long distances to sparsely populated areas (Yousefi Mojir & Pilemalm, 2016). The occurrence of the COVID-19 pandemic also showed how these challenges could be exacerbated under extreme situations. Shortage of resources and longer travel distances translate to longer response times, especially for rural and sparsely populated areas. Longer response times can lead to higher rate of damage to the properties and the environment, harm to human health, and higher mortality rates, for example in cases of cardiac arrest (Waalewijn et al., 2001) and fires (Jaldell, 2004 & 2017). Consequently, emergency services in recent years have tried to use additional resources that will not incur additional cost but can help them improve their response times. One such resource is civilians, who volunteer to help in case of emergencies in their vicinity.

In collaboration with the municipal fire and rescue services (FRS), volunteers in some villages in Sweden, where the emergency coverage is poor, are trained (and potentially equipped) for first response to a selected set of emergency types. The civilians in this initiative receive training in cardiopulmonary resuscitation (CPR), first aid, extinguishing of small fires, and required actions when a person first arrives at some emergencies (e.g., a fire or a traffic accident) (Hallström et al., 2017). These trained civilians are, simultaneous with the dispatch of emergency services, notified of an emergency through a short message service (SMS) or an application on their mobile phones. Training and potentially equipping civilians in rural areas requires monetary and time investment from the emergency services. However, as the FRS have budget (and personnel) restrictions, it is not possible to include a large number of civilians and villages in such a volunteer initiative, especially not in the first round. While a major part of the cost is setup, there are also running costs for maintaining the volunteer centers, including costs for continuous training, and possibly equipment. It is therefore important to initially select a number of

villages that both meet a few minimum requirements for inclusion and gives the best improvement for the emergency response system. However, it is not always trivial to determine which villages will give the most benefit. A possible solution for this problem is to use optimization and develop a model that can take into account the required factors. However, such a model also needs to be comprehensible for the practitioners, which means that a simpler model that can fulfill the purpose might be more suitable than a very complex one.

The aim of the paper is to develop an optimization model to determine which rural villages should be included in a volunteer first response initiative, so that the population in the region receive the best possible emergency coverage from FRS and the volunteer initiative.

The rest of this paper is organized as follows. In the next section, we provide a brief overview of the related literature and positioning of this work with respect to the literature. In the following section, we present the problem description as well as the proposed optimization model. Then, we describe the numerical experiments, detailing the case study, where we help an FRS organization to select rural villages to use as volunteer centers, and the results. Finally, in the last section, we present conclusions and future research directions.

PREVIOUS STUDIES

Emergency response coverage in rural areas in many countries, including Sweden, is weak. As Pilemalm et al. (2013) pointed out, most of Sweden can be considered rural, given that population centers are sparsely located and driving time (by a car) to the nearest town with more than 3000 residents takes more than 45 minutes from many areas. Therefore, strengthening the emergency response for rural areas becomes even more crucial. In the following, we briefly discuss some of the literature that are most relevant to this study.

Volunteer initiatives have emerged and increased in numbers over a relatively short amount of time as a potential solution to enhancing emergency response time and coverage, especially in rural areas. In some initiatives, volunteers respond to only a specific type of emergency (e.g., out-of-hospital cardiac arrest - OHCA). Examples include SMS lifesavers in Sweden (https://heartrunner.com/) and HartslagNu in the Netherlands (https://hartslagnu.nl/), but more examples can be found in several reviews such as Oving et al. (2019), Scquizzato et al. (2020), and Valeriano et al. (2021). In other initiatives, a wider range of emergency types is considered; examples are the enhanced neighbors project (Ramsell et al., 2017) and civilian response persons (Stenberg et al., 2014; Hallström et al., 2017) in Sweden. Here, volunteers respond to selected medical emergencies as well as small fires and traffic accidents. There are also initiatives in Germany for responding to medical emergencies (e.g., Gross et al. (2019)) and in Finland (Tamminen et al., 2019). Several studies, especially within the medical area, have examined the impact of volunteer responses on the outcome of the cases, showing that volunteers can reduce response times (see e.g., Auricchio et al. (2019), Berglund et al. (2018), Ringh et al. (2015), and Stroop et al. (2020)). Svensson et al. (2024) compared emergency service response times between rural and urban areas, showing a notable difference. They also showed that volunteers had the shortest median response times to OHCA compared to emergency medical services (EMS) and FRS. Similarly, Lapidus et al. (2023) showed that volunteers had shorter response times than EMS, especially in sparsely populated areas.

Two crucial aspects for being able to start and maintain volunteer initiatives, are to attract and motivate volunteers. Thus, it is important to understand which factors lead to the involvement of civilians in volunteer initiatives and why they might leave. Cowlishaw et al. (2008) raised family issues as a factor with direct impact on the decrease of the number of volunteers in rural areas. Timmons and Vernon-Evans (2013) as well as Barry et al. (2019) showed that one of the motivating factors for people to volunteer comes from a personal level, their “ideas of altruism and sense of community” and personal experiences of emergencies or illness. Choudhury (2010) stated that local government agencies are important bodies for attracting and managing volunteers.

In many volunteer initiatives, volunteers are alerted through their mobile phones, for example by SMS or specific mobile phone applications (Matinrad & Granberg, 2023). Initiatives such as the enhanced neighbors project relied on SMS for sending notifications to volunteers (Ramsell et al., 2017). SMS lifesavers also started using SMS, but later switched to a mobile phone application (Jonsson et al., 2020), while other initiatives, HartslagNu for example, use both SMS and a mobile phone application (Slaa, 2020). All these initiatives use some type of geographical positioning of the volunteers so that they can send the notification to the volunteers in the vicinity of the incident site. This geographical positioning can be static (i.e., registered addresses) or dynamic using Global Navigation Satellite Systems (GNSS). All these initiatives need some type of decision-making logic, which can be embedded in a decision support system (DSS). The underlying logic in an emergency DSS can be a simple rule such as the one used in SMS lifesavers initiative (Granberg et al., 2022) or it can be more advanced models (see e.g., Matinrad and Granberg (2023) and Frigstad et al. (2023)).

For rural areas, determining which villages that should be included in a volunteer first response initiative relies

on different factors, where one crucial factor is the geographical distribution of the villages. While the literature on volunteer location studies is scarce, there are many studies to be found on the location of emergency service stations (see for example, van den Berg and Aardal (2015), Røislien et al. (2016), van den Berg et al. (2016), Garner and van den Berg (2017), and Jagtenberg and Mason (2020)) or public emergency equipment such as automated external defibrillators (AEDs) (e.g., Chan et al. (2013), Bonnet et al. (2015), and Yang et al. (2021)). The majority of these studies use a type of covering location model, such as the maximal coverage location problem (MCLP) (Church & ReVelle, 1974) or the maximum expected coverage location problem (MEXCLP) (Daskin, 1983). In some studies, such as Bonnet et al. (2015), a MCLP model is extended to incorporate more factors, including temporal aspects (e.g., accessibility of candidate locations across different times for AED locations) and amount of used resources (e.g., number of deployed devices).

In this paper, we address the problem of determining which rural villages that should serve as volunteer centers and present an optimization model that aims to maximize the population coverage, with regards to both volunteer and FRS response. To the best of our knowledge, this is the first optimization model for such a problem. Given the simplicity of the model, it is easily comprehensible for practitioners. Results from the model are currently being utilized in an ongoing volunteer initiative in Sweden for creating volunteer centers across rural areas in order to improve the emergency response system.

PROBLEM DESCRIPTION AND MATHEMATICAL MODEL

In this section, we will first briefly describe the problem statement as well as assumptions that we have made, and then, we will present the optimization model.

Problem Statement and Assumptions

The problem, Selection Of Rural Villages as Volunteer centers for Emergency response (SORVIVE), deals with *determining which rural villages should be included in the volunteer response initiative, aiming to improve the emergency response coverage, considering the current professional emergency service system*.

In this problem, we have made a few assumptions:

- The candidate villages are known and have at least five people that are willing to volunteer and respond to emergencies if they are notified. It is hypothetically possible to consider all villages as candidates, but since one person will not always be able to respond, more than one volunteer is needed to get any effect. Furthermore, in these kinds of initiatives, the volunteers prefer not to be the only one responding to an incident (Ramsell et al., 2017). Thus, it is preferable that the candidate villages have at least five residents willing to volunteer.
- Volunteers in the selected villages are assumed to primarily respond to emergencies within their own villages. However, if their neighboring villages are within a reasonable distance, they can receive help as well.

In Figure 1, we show a schematic representation of the SORVIVE problem, with two FRS stations in the area, and four selected villages that will function as volunteer centers. Note that in this schematic representation, we have considered different coverage radius for different villages, given the size of the villages and their geographical position, and a 20-minute coverage radius for the FRS stations. Each of these volunteer centers will cover a subset of villages and towns (population centers) that can be divided into population squares.

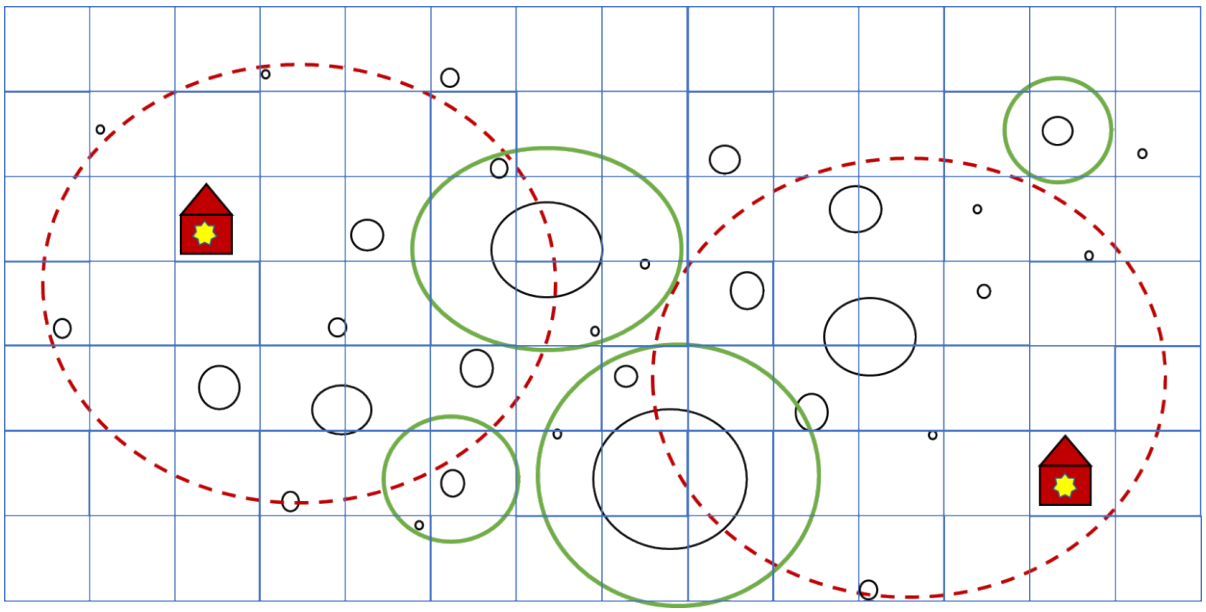


Figure 1. Schematic representation of the problem.
The small black circles represent rural villages, and the blue squares represent population squares; the dashed red circles show the coverage radius of FRS with the red houses representing the FRS stations, and the solid green circles show the coverage area of selected population squares anchored around a village.

Model Formulation

As input to the mathematical formulation of the SORVIVE problem, we know where the FRS stations as well as candidate volunteer centers across a specific region are located. We can also divide the whole region into squares where we know the resident population in each square, and calculate all distances between stations, centers, and squares. We also know the coverage radius for volunteer centers (either individually based on their size and location or a common radius for all of them) and the maximum service time to reach a population center from an FRS station. It is also known how many volunteer centers are planned to be established. The model determines which candidate volunteer centers that are selected, and which population squares they cover.

We have used the following notation in the mathematical formulation of the SORVIVE problem:

Sets and indices

$I = \{1, \dots, m\}$	Set of candidate volunteer centers, indexed by $i \in I$
$K = \{1, \dots, o\}$	Set of population squares, indexed by $k \in K$
$J = \{1, \dots, n\}$	Set of FRS stations, indexed by $j \in J$

Parameters

s_k	Total number of people in population square *k*
z_{jk}	Equals to 1 if FRS station *j* covers population square *k*; otherwise, 0
df_{kj}	(Shortest) response time (i.e., preparation time plus travel time) between FRS station *j* and population square *k*
dp_{ik}	(Shortest) response time between volunteer center *i* and population square *k*
P	Maximum number of volunteer centers
D_i	Coverage radius of volunteer center *i*
L_j	A predefined maximum service time to reach a region from FRS station *j* within a given time

Decision variables

y_i	Equals to 1 if volunteer center is established in *i*; otherwise, 0
x_k	Equals to 1 if population square *k* is covered; otherwise, 0
w_{ik}	Equals to 1 if volunteer center at village *i* covers population square *k*; otherwise, 0

The SORVIVE problem is thus formulated as an integer programing (IP) model:

$$max \sum_{k \in K} s_k x_k \quad (1)$$

$$\sum_{i \in I} y_i \leq P \quad (2)$$

$$D_i y_i \geq dp_{ik} w_{ik} \qquad \forall i \in I, k \in K \quad (3)$$

$$\sum_{i \in I} w_{ik} \leq 1 \qquad \forall k \in K \quad (4)$$

$$x_k \leq \sum_{i \in I} w_{ik} + \sum_{j \in J} z_{jk} \qquad \forall k \in K \quad (5)$$

$$y_i, w_{ik}, x_k \in \{0,1\} \qquad \forall i \in I, j \in J, k \in K \quad (6)$$

The objective function (1) maximizes the weighted coverage of the population squares, where the weight is the population. Constraint (2) is defined to limit the number of the selected volunteer centers to *P*. Constraint set (3) determines which population squares a selected volunteer center covers; coverage is possible if the travel time between a center and a square is shorter than the predefined coverage radius for the volunteer center. Constraints (4) ensure that each square is covered by a maximum of one volunteer center. Constraint set (5) determines whether a square is covered or not, where the coverage can be by FRS stations, by a volunteer center, or by both FRS stations and a volunteer center. Constraints (6) are the binary constraints for the variables.

NUMERICAL EXPERIMENTS

In the following, we first present a short description of the case study that we used to test and validate the model and then, we present the numerical results.

Case Description

We used the area of a large FRS organization (Räddningstjänsten Östra Götaland, RTÖG) in Östergötland county in Sweden as our use case. Östergötland is located at the south-east of Sweden with a population of over 470 000 inhabitants and an area of 9 979 square kilometers, leading to a population density of approximately 47.1 inhabitants per square kilometers. Östergötland is divided into 13 municipalities, and RTÖG is responsible to provide FRS in five of these, that together have about 341 140 (72.6% of the county total) inhabitants and spans an area of 5 270 square kilometers (53% of the county). In RTÖG's area of responsibility, there are a total of 120 populations centers, including cities, towns, and villages. A subset of these (the ones where we expect to be able to recruit at least five volunteers) are used as potential volunteer centers. The area is also divided into a set of 6084 population squares, with a known number of permanent residents. This is done to get a detailed view of the expected coverage from fire stations and volunteers centers.

RTÖG currently has 20 fire stations: four fulltime stations, 12 part-time stations, and four volunteer stations. At the fulltime stations, there is always at least one unit (typically consisting of four firefighters and one commander) present. They have a preparation time (i.e., the time from when the alert comes to the station until the resources leave the station) of 90 seconds. At the part-time stations, the firefighters have preparedness at home or at their regular job. When there is an alert, they will go to the station to prepare and leave in an appropriate vehicle. Thus, their preparation time is longer than for the fulltime firefighters, being 300 seconds (five minutes). Volunteer firefighters do not have preparedness and are thus not obligated to respond when there is an alert (although many will do so). Otherwise, a volunteer station is much like a part-time station. There is no set preparation time, since response is voluntary, but for calculation purposes, we use 300 seconds. Volunteer firefighters, and volunteer FRS stations, are different from the civilian volunteer centers that we locate in this case study. While the volunteer firefighters have training, clothes, vehicles and a fire station to which they go before responding, the civilian first response volunteers that are to be trained at the volunteer centers, will have much more limited training and equipment, and will typically respond directly from where they are when they get the alert.

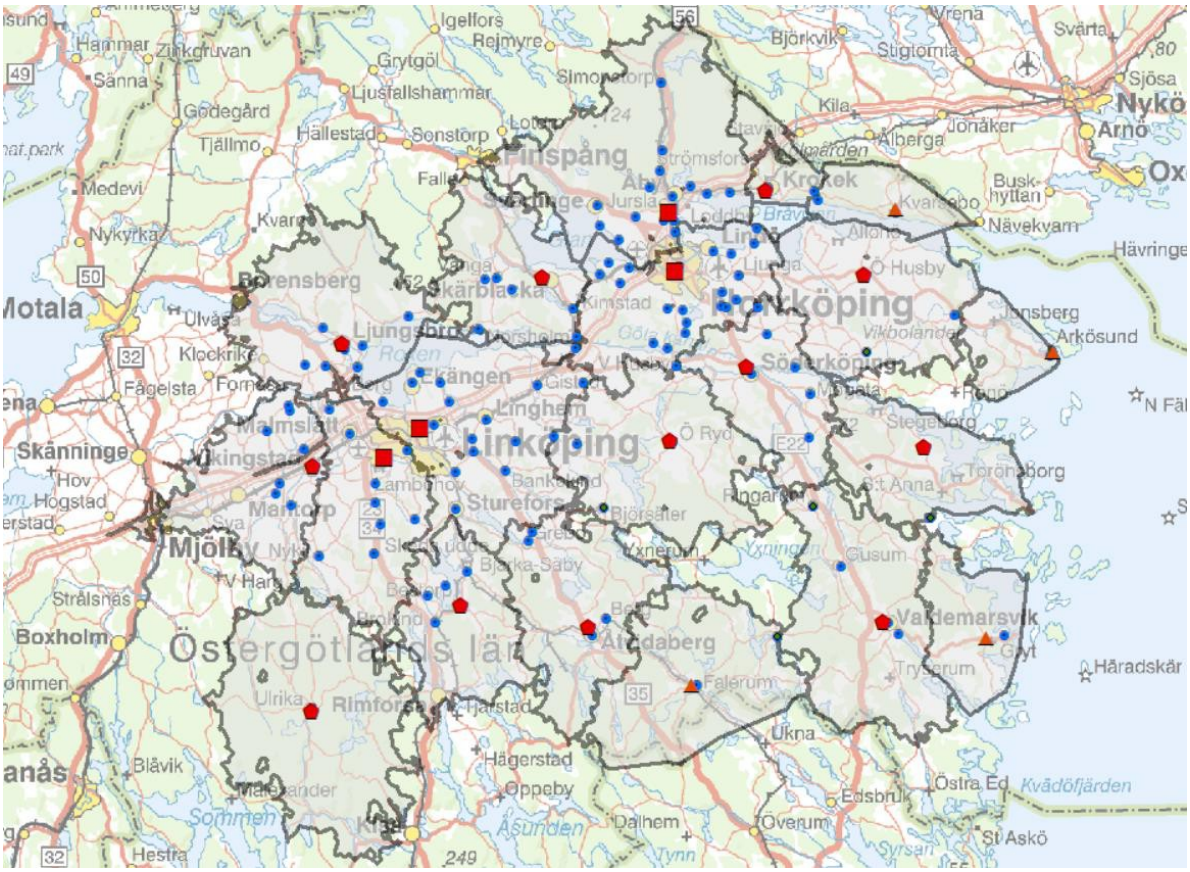
RTÖG plans to start some volunteer centers across the county, and as their starting goal, they plan to begin with initiating two centers. Thus, our initial target for locating volunteer centers is $P = 2$. As previously indicated, the

location of all population centers as well as FRS stations are known, and thus, travel times between each pair of these locations can be easily calculated. We used ArcGIS 10.6.1, utilizing the Network Analyst module to calculate these, based on the Swedish national road database, and using shortest path calculations with the speed limit and link length to estimate the time it takes to travel on a link.

We considered the service radius for all FRS stations (fulltime, part-time, and volunteer) to be 1200 seconds (20 minutes) and for the volunteer centers to be 600 seconds (10 minutes). The preparation time for the fulltime stations was set to 90 seconds, while this time for part-time and volunteer FRS stations was set to 300 seconds.

We considered two scenarios where volunteers had either 90 seconds or 300 seconds preparation time, since real preparation times for volunteers are difficult to estimate. In one additional setting, as a verification, we excluded the volunteer FRS stations from the set of FRS stations, which resulted in a total of 16 stations (rather than the original 20).

In Figure 2, we show the primary coverage areas for the 20 FRS stations, where the border is reached within 20 minutes of driving from a station; the volunteer FRS stations are shown with the red triangles and the remaining are fulltime and part-time FRS stations.


Hammar
Zinkgruvan
Grytgöl
Ljusfallshammar
Sånna
Gödegård
Tjällmo
Hällestad
Sonstorp
Rejmyre
Igelfors
Björkvik
Vrena
Svärta
Sjösa
Kila
Jonåker
Arnö
Nyköping
Alberga
Finspång
Medevi
Nykyrka
Kvarn
Borensberg
Motala
Ulvåsa
Klockrike
Fågelsta
Malmslätt
Linköping
Skänninge
Hov
Högstad
Mjölby
Strålsnäs
Boxholm
Östergötlands län
Åtvidaberg
Valdemarsvik
Gryt
Häradskär
Tryserum
Falerum
Ukna
Hägerstad
Oppeby
Blåvik
Edsbruk
Östra Ed
Kvädöfjärden
Dalhem
Överum
St Askö
Hestra
Åsunden
Sommen
Arkösund
Hävringe
Nävekvarn
Busk-hyttan
Ljungsbro
Linghem
Stegeborg
Torpsborg
Vikbolandet
Ingarö
Husby
Ljuda
Kimstad
Skärblacka
Rimforsa
Ulrika
Vrena
Bankeryd
Sturefors
Malmslätt
Kvarsebo
Krokek
Jursla
Fall
Lotorp
Simonstorp
Igelfors
Rejmyre
Grytgöl
Ljusfallshammar
Sånna
Gödegård
Tjällmo
Hällestad
Sonstorp
Kila
Jonåker
Alberga
Arnö
Nyköping
Svärta
Sjösa
Björkvik
Stigtorp
Vrena
Medevi
Nykyrka
Kvarn
Borensberg
Motala
Ulvåsa
Klockrike
Fågelsta
Malmslätt
Linköping
Skänninge
Hov
Högstad
Mjölby
Strålsnäs
Boxholm
Östergötlands län
Åtvidaberg
Valdemarsvik
Gryt
Häradskär
Tryserum
Falerum
Ukna
Hägerstad
Oppeby
Blåvik
Edsbruk
Östra Ed
Kvädöfjärden
Dalhem
Överum
St Askö
Hestra
Åsunden
Sommen
Arkösund
Hävringe
Nävekvarn
Busk-hyttan
Ljungsbro
Linghem
Stegeborg
Torpsborg
Vikbolandet
Ingarö
Husby
Ljuda
Kimstad
Skärblacka
Rimforsa
Ulrika

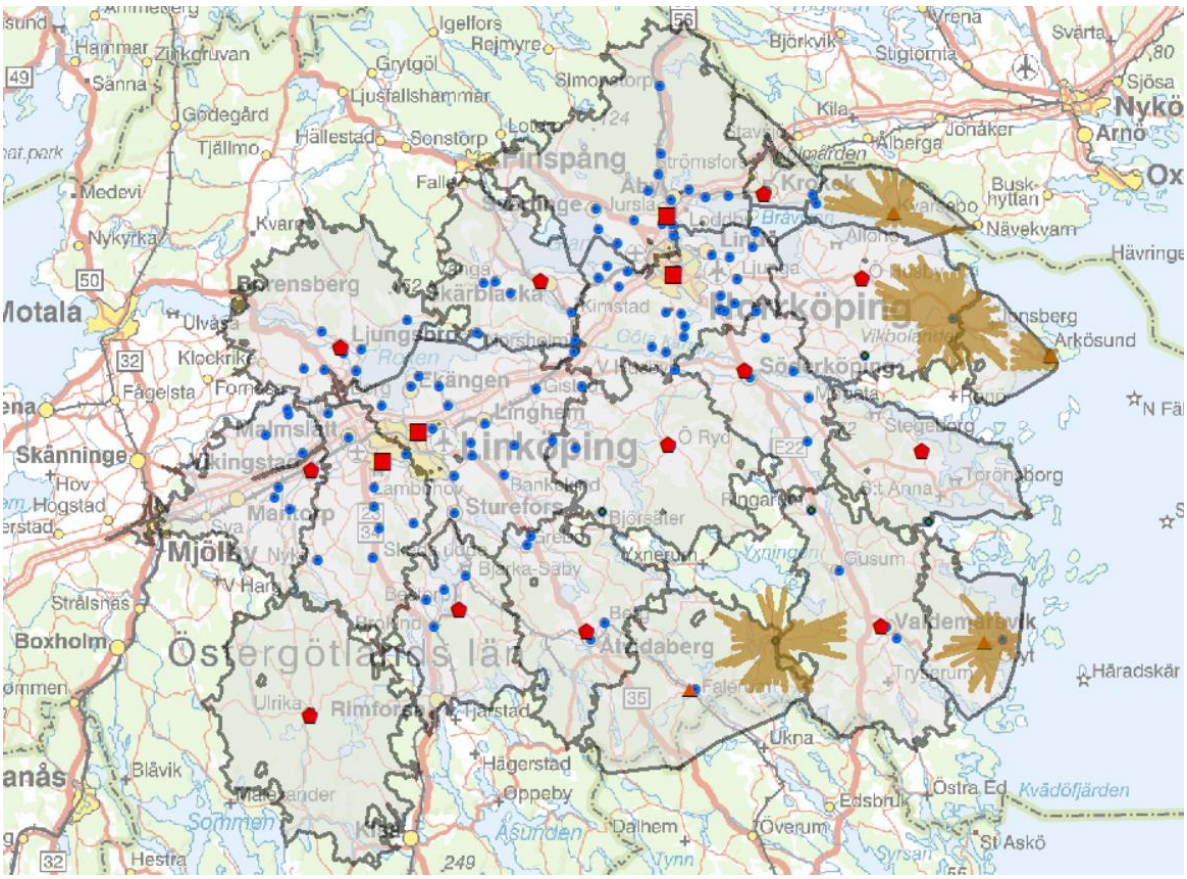


Figure 3. Location of five volunteer centers with 90 seconds preparation time, not considering volunteer FRS stations

Then, we wanted to compare output of the model when including all stations (20 stations) and investigate how different preparation times might affect the coverage. In Table 1, we present results for $P = 2$ but for different setups where volunteers had 90 seconds or 300 seconds as their preparation time. For both setups, the model produces the same optimal locations, indicating stability of the model in producing the results, that is, the variability in preparation time does not affect which centers that are opened, and has only a marginal impact on the coverage. As expected, with the longer volunteer preparation time, a slight decrease in the total coverage rate can be seen as the response time (preparation time plus travel time to the scene) becomes longer. The decrease can, however, be considered fairly small.

Table 1. Comparison of coverage for different data setups, $P = 2$

Row	Data setup	Total number of covered people	Total coverage
1	90 seconds, 20 stations	335 381	98.3%
2	300 seconds, 20 stations	335 293	98.3%

Finally, we ran the model with all 20 FRS stations included, volunteers had 90 seconds preparation time (and 10 minutes coverage radius), and for different values of P. The result of this test for $P = 5$ is visualized in Figure 4. The solid, red circle (Ringarum) shows one of the first selected volunteer centers by the model (i.e., when the model was run for $P = 2$); according to RTÖG, the FRS used to have a part-time fire and rescue service station there, that was closed down in recent years. However, as we can see, parts of the Ringarum area receives a poor coverage from the closest FRS stations, which was also confirmed by RTÖG. Consequently, locating a volunteer center in that region, as shown in Figure 4, was considered reasonable by RTÖG. The dashed, blue circle in Figure 4 is around Ramsdal, the second village which was selected by RTÖG to be a volunteer center, and the second village selected by the model for $P = 2$. Ringarum had a population of 606 residents in 2020 and Ramsdal had 50, but together they cover a total of 3003 inhabitants within 10 minutes. After the selection of these two villages, RTÖG started recruitment and training of volunteers in these two locations.

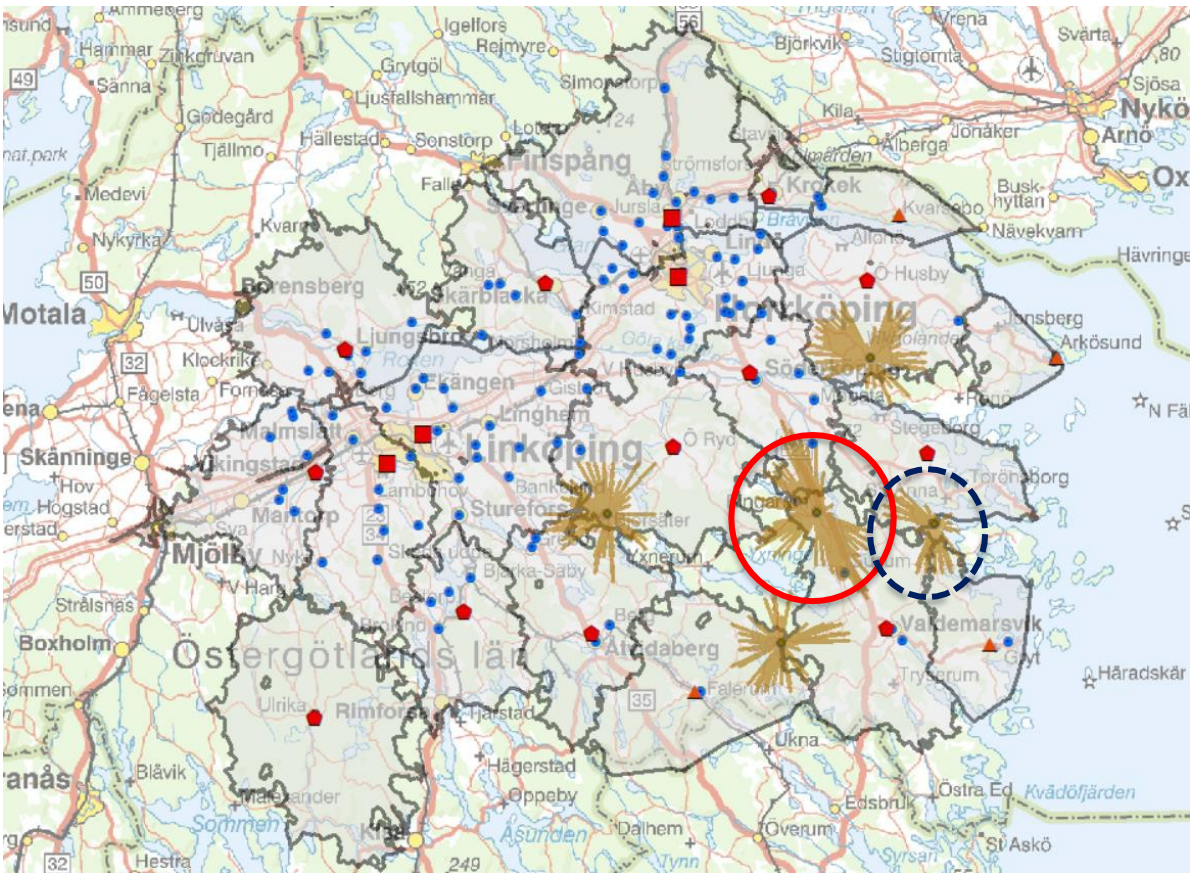


Figure 4. Location of five volunteer centers with 90 seconds preparation time, considering all 20 FRS stations

CONCLUSIONS AND FUTURE STUDIES

In this work, we proposed a simple coverage model for the selection of rural villages as volunteer centers to improve emergency response coverage. The model showed promising results, which were acceptable for practitioners (i.e., the largest FRS organization in Östergötland), and the simplicity of the model makes it easy for them to understand and utilize it. Thus, they picked a solution from the model, and have started volunteer centers in two new villages in Östergötland. This shows that the model is useful as a practical decision support tool. We are now following the implementation process, and when the volunteer initiative has matured, and the volunteers in these two villages have responded to enough events, we will evaluate the practical effects of the initiative. This should give valuable insight into how information system facilitated volunteerism, combined with operations research-based techniques can help improve the emergency response system.

The optimization model can be further extended by considering additional aspects, such as minimizing response times, incorporating uncertain elements, and including budget constraints. However, this must be done carefully, so that the comprehensibility of the model is not lost. Moreover, further research directions related to the volunteer initiative can include studies aimed at assessing the effectiveness of volunteer response and other aspects that should or could be incorporated.

ACKNOWLEDGMENTS

We kindly thank Michael Wahlqvist, Marie Bäck, Joakim Ivarsson, and the rest of the staff involved in the volunteer initiation project at RTÖG for kindly sharing knowledge and data. We are also grateful to the track chair responsible for handling our paper as well as the anonymous reviewers for their valuable comments. This work was funded by the Swedish civil contingencies agency (MSB), through the research program Managing the incident site of the future (MIST), which is managed by the Center for advanced research in emergency response (CARER).

REFERENCES

Auricchio, A., Gianquintieri, L., Burkart, R., Benvenuti, C., Muschietti, S., Peluso, S., Mira, A., Moccetti, T., & Caputo, M. L. (2019). Real-life time and distance covered by lay first responders alerted by means of smartphone-application: Implications for early initiation of cardiopulmonary resuscitation and access to automatic external defibrillators. *Resuscitation*, *141*(December 2018), 182–187. https://doi.org/10.1016/j.resuscitation.2019.05.023

Barry, T., Guerin, S., & Bury, G. (2019). Motivation, challenges and realities of volunteer community cardiac arrest response: a qualitative study of 'lay'community first responders. *BMJ Open*, *9*(8), e029015.

Berglund, E., Claesson, A., Nordberg, P., Djärv, T., Lundgren, P., Folke, F., Forsberg, S., Riva, G., & Ringh, M. (2018). A smartphone application for dispatch of lay responders to out-of-hospital cardiac arrests. *Resuscitation*, *126*, 160-165.

Bonnet, B., Dessavre, D. G., Kraus, K., & Ramirez-Marquez, J. E. (2015). Optimal placement of public-access aeds in urban environments. *Computers & Industrial Engineering*, 90, 269–280.

Chan, T. C., Li, H., Lebovic, G., Tang, S. K., Chan, J. Y., Cheng, H. C., Morrison, L. J., & Brooks, S. C. (2013). Identifying locations for public access defibrillators using mathematical optimization. *Circulation*, 127(17), 1801–1809.

Choudhury, E. (2010). Attracting and managing volunteers in local government. *Journal of Management Development*, *29*(6), 592–603. https://doi.org/10.1108/02621711011046558

Church, R., & ReVelle, C. (1974, December). The maximal covering location problem. In *Papers of the regional science association* (Vol. 32, No. 1, pp. 101-118). Berlin/Heidelberg: Springer-Verlag.

Cowlishaw, S., Evans, L., & McLennan, J. (2008). Families of rural volunteer firefighters. *Rural Society*, *18*(1), 17–25. https://doi.org/10.5172/rsj.351.18.1.17

Daskin, M. S. (1983). A maximum expected covering location model: formulation, properties and heuristic solution. *Transportation science*, *17*(1), 48-70.

Frigstad, L., Furu, V., Svenkerud, S. K., Claesson, A., Andersson, H., & Granberg, T. A. (2023). Joint planning of drones and volunteers in emergency response to out-of-hospital cardiac arrest. *Computers & Industrial Engineering*, *185*, *109648*.

Garner, A. A., & van den Berg, P. L. (2017). Locating helicopter emergency medical service bases to optimise population coverage versus average response time. *BMC emergency medicine*, *17*(1), 1-11.

Granberg, T. A., Erlander, S., Fredman, D., Olovsson, L., & Persson, E. (2022). Predicting Volunteer Travel Time to Emergencies. In *19th International Conference on Information Systems for Crisis Response and Management, Tarbes, France, May 22-25, 2022* (pp. 44-54).

Gross, B., Schanderl, F., Staedt, N., & Elsner, C. (2019). Dispatching first responders via mobile application systems: Status quo of implementation in Germany, evidence and challenges during introduction. *Notfall+ Rettungsmedizin*, *22*, 483-491.

Hallström, V., Ramsell, E., & Yousefi Mojir, K. (2017). Organiserade räddningsinsatser med frivilliga. Presentation av några verksamheter som initierades i Sverige mellan 2010-2016 [Organized rescue efforts with volunteers. Presentation of some activities that were initiated in Sweden between 2010-2016]. *Swedish Civil Contingencies Agency, Linköping*.

Jagtenberg, C. J., & Mason, A. J. (2020). Improving fairness in ambulance planning by time sharing. *European Journal of Operational Research*, *280*(3), 1095-1107.

Jaldell, H. 2004. *Tidsfaktorns betydelse vid räddningsinsatser*.

Jaldell, H. (2017). How important is the time factor? Saving lives using fire and rescue services. *Fire technology*, *53*(2), 695-708.

Jonsson, M., Berglund, E., Djärv, T., Nordberg, P., Claesson, A., Forsberg, S., Nord, A., Tan, H. L., & Ringh, M. (2020). A brisk walk—Real-life travelling speed of lay responders in out-of-hospital cardiac arrest. *Resuscitation*, *151*, 197–204.

Lapidus, O., Jonsson, M., Svensson, L., Hollenberg, J., Berglund, E., Riva, G., Claesson, A., Nordberg, P., Rosenqvist, M., Forsberg, S., Nord, A., & Ringh, M. (2023). Effects of a volunteer responder system for out-

of-hospital cardiac arrest in areas of different population density–A retrospective cohort study. *Resuscitation*, *191*, 109921.

Matinrad, N. (2022). *Models for Dispatch of Volunteers in Daily Emergency Response* (Doctoral dissertation, Linköping University Electronic Press).

Matinrad, N., & Granberg, T. A. (2023). Optimal pre-dispatch task assignment of volunteers in daily emergency response. *Socio-Economic Planning Sciences*, *87*, 101589.

Oving, I., Masterson, S., Tjelmeland, I. B. M., Jonsson, M., Semeraro, F., Ringh, M., Truhlar, A., Cimpoesu, D., Folke, F., Beesems, S. G., & ESCAPE-NET Investigators. (2019). First-response treatment after out-of-hospital cardiac arrest: a survey of current practices across 29 countries in Europe. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, *27*(1), 112.

Pilemalm, S., Stenberg, R., & Granberg, T. A. (2013). Emergency Response in Rural Areas. *International Journal of Information Systems for Crisis Response and Management*, *5*(2), 19–31. https://doi.org/10.4018/jiscrm.2013040102

Quarantelli, E. L. (1995). Disasters Are Different, Therefore Planning For And Managing Them Requires Innovative As Well As Traditional Behaviors. *Disasters*.

Ramsell, E., Pilemalm, S., & Granberg, T. A. (2017). Using Volunteers for Emergency Response in Rural Areas – Network Collaboration Factors and IT support in the Case of Enhanced Neighbors. *14th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2017)*, 985–995.

Ringh, M., Rosenqvist, M., Hollenberg, J., Jonsson, M., Fredman, D., Nordberg, P., Järnbert-Pettersson, H., Hasselqvist-Ax, I., Riva, G., & Svensson, L. (2015). Mobile-Phone Dispatch of Laypersons for CPR in Out-of-Hospital Cardiac Arrest. *New England Journal of Medicine*, *372*(24), 2316–2325. https://doi.org/10.1056/NEJMoa1406038

Røislien, J., van den Berg, P. L., Lindner, T., Zakariassen, E., Aardal, K., & van Essen, J. T. (2016). Exploring optimal air ambulance base locations in Norway using advanced mathematical modelling. *Injury prevention*.

Scquizzato, T., Pallanch, O., Belletti, A., Frontera, A., Cabrini, L., Zangrillo, A., & Landoni, G. (2020). Enhancing citizens response to out-of-hospital cardiac arrest: A systematic review of mobile-phone systems to alert citizens as first responders. *Resuscitation*, *152*(May), 16–25. https://doi.org/10.1016/j.resuscitation.2020.05.006

Slaa, G. (2020). *Increasing cardiac arrest survival by improving the volunteer alerting algorithm* (Master's thesis, University of Twente).

Stenberg, R., Pilemalm, S., & Yousefi Mojir, K. (2014). *Samverkansformer för räddning och respons Exempel och förslag [Forms of collaboration for rescue and response Examples and suggestions]* (Issue 12). https://liu.se/artikel/carer-rapportserie

Stroop, R., Kerner, T., Strickmann, B., & Hensel, M. (2020). Mobile phone-based alerting of CPR-trained volunteers simultaneously with the ambulance can reduce the resuscitation-free interval and improve outcome after out-of-hospital cardiac arrest: A German, population-based cohort study. *Resuscitation*, *147*(March 2019), 57–64. https://doi.org/10.1016/j.resuscitation.2019.12.012

Svensson, A., Nilsson, B., Lantz, E., Bremer, A., Årestedt, K., & Israelsson, J. (2024). Response times in rural areas for emergency medical services, fire and rescue services and voluntary first responders during out-of-hospital cardiac arrests. *Resuscitation plus*, *17*, 100548.

Tamminen, J. I., Hoppu, S. E., & Kämäräinen, A. J. (2019). Professional firefighter and trained volunteer first-responding units in emergency medical service. *Acta Anaesthesiologica Scandinavica*, *63*(1), 111-116. Timmons, S., & Vernon-Evans, A. (2013). Why do people volunteer for community first responder groups? *Emergency Medicine Journal*, *30*(3), 1–5. https://doi.org/10.1136/emermed-2011-200990

Valeriano, A., Van Heer, S., de Champlain, F., & Brooks, S. C. (2021). Crowdsourcing to save lives: A scoping review of bystander alert technologies for out-of-hospital cardiac arrest. *Resuscitation*, *158*(July), 94–121. https://doi.org/10.1016/j.resuscitation.2020.10.035

van den Berg, P. L., & Aardal, K. (2015). Time-dependent MEXCLP with start-up and relocation cost. *European Journal of Operational Research*, *242*(2), 383-389.

van den Berg, P. L., Kommer, G. J., & Zuzáková, B. (2016). Linear formulation for the maximum expected coverage location model with fractional coverage. *Operations Research for Health Care*, *8*, 33-41.

Waalewijn, R. A., de Vos, R., Tijssen, J. G., & Koster, R. W. (2001). Survival models for out-of-hospital cardiopulmonary resuscitation from the perspectives of the bystander, the first responder, and the paramedic. *Resuscitation*, *51*(2), 113-122.

Yang, L., Xiong, L., & Yang, W. (2021). An accessibility spatial search algorithm to optimize defibrillator deployment in indoor space. *International Journal of Geographical Information Science*, 35(6), 1251–1272.

Yousefi Mojir, K., & Pilemalm, S. (2016). Actor-centred emergency response systems: a framework for needs analysis and information systems development. *International Journal of Emergency Management*, *12*(4), 403–434. https://doi.org/10.1504/IJEM.2016.079844

Civilian-Military Integration in Ukrainian Defense Supply Chain

Myroslava Kushnir
Ukrainian Catholic University
makushnir@ucu.edu.ua

Anna Nagurney
Isenberg School of Management,
University of Massachusetts Amherst
nagurney@isenberg.umass.edu

Renata Konrad
The Business School
Worcester Polytechnic Institute
rkonrad@wpi.edu

ABSTRACT

This research delves into the intricacies of the Ukrainian military's supply chain during the ongoing Russian-Ukrainian war, with a specific focus on the significant role played by the civilian population in supporting the armed forces. Termed a defense supply chain, our study employs a multilevel model encompassing logistical, informational, and financial dimensions to elucidate the multifaceted levels within the defense supply chain network. Key actors identified include manufacturers, non-governmental organizations (NGOs), donors, and conscripts. The study sheds light on the vulnerabilities in the flow of goods in the defense supply chain. A resilient defense supply chain relies on efficient NGO communication and the crucial alignment of donor income growth with escalating resource prices.

Keywords

Civil-Military Integration, Defense Supply Chain, Russian-Ukrainian War, Supply Chain Network Model.

INTRODUCTION

The year 2022 marked a pivotal moment in global affairs as Russia's full-scale invasion of Ukraine reshaped the geopolitical landscape, prompting a profound reevaluation of military capabilities, strategic responses, and international alliances. Against this backdrop, our research characterizes the intricate web of dynamic civilian-military interactions within the context of the Ukrainian defense supply chain and highlights the role of information in this context.

The Russian-Ukrainian war, characterized by a full-scale invasion, unfolded as an asymmetric warfare scenario where Russia possesses substantial resources over Ukraine (International Institute for Strategic Studies, 2021). This asymmetry catalyzed a distinct military campaign heavily reliant on the self-organization of civilians. Ukrainian self-organization, which, as noted by Channel-Justice (2022), was already evident over a decade ago, as the 2013-2014 Euromaidan protests were unfolding in Ukraine.

Ukraine's two-year defensive challenges initial perceptions regarding the military power balance between Russia and Ukraine (Dalsjö et al, 2023; Ti, 2023). Increasing evidence points to the aggressor's ill-judged and overly optimistic military calculations such as Russia's poor understanding of its own forces, poor Russian training, and low standards of tactical competence. A key, yet understudied, element in Ukraine's surprising success was the role of its citizens. Indeed, Barany (2023), Hackett (2023), Hoffman (2022), and Tierney (2023) underscore the significance of the human factor in warfare, emphasizing the importance of investing in personnel, competent commanders, and comprehensive training.

Armed conflict triggers the application of the Third Geneva Convention of 1949, the First Additional Protocol of 1977, and customary rules of International Humanitarian Law (IHL). Humanitarian organizations, guided by IHL principles, maintain a clear demarcation between military and nonmilitary aid, underscoring the imperative of

CoRe Paper – Usability and Universal Design of ICT for Emergency Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

principled neutrality to protect civilians and to uphold human rights (Melzer, 2019).
Ley (2022) emphasizes the fundamental role of social networks in mobilizing against insecurity, transforming individual fears into collective anger and providing a sense of security. For Ukrainian civilians, providing aid to the military is viewed as a humanitarian response, as it has helped to stave off the Russian invasion and the resulting horrific damage and suffering to the population (Dunn et al., 2023). Thus, volunteer action was overtly political, a form of resistance to Russian attacks and domination (Konrad et al., 2023; Slim, 2022).

We define the current situation in Ukraine as a full-scale invasion with a defense supply chain. By full-scale invasion we mean that one government's armed forces are defending their sovereign country against an unprovoked invasion by armed forces of another state. The term defense supply chain refers to a logistical system in which civilians are an essential and integral part for a defending nation's military success. Specifically, we study how civilians provide the Ukrainian military with ammunition, food, transport, medical supplies, warehouses, and act as the main channels for information transmission. While civilian surveillance, and participatory gathering and sharing of intelligence have a growing body of evidence, the agentic power of civilians in military logistics has not been documented. Although defense supply chain network economics has been studied (Nagurney, 2023), it has been examined in the context of behavior of defense firms, which are concerned about revenues as well as risk. Moreover, the role of information in defense supply chains has not received much attention and, yet, as we hypothesize, information plays an indispensable role in supply chain effectiveness, by providing the crucial linkages between the logistical and financial domains.

Motivated by our observations in Ukraine, our research question is: to what extent is the defense supply system in Ukraine vulnerable, and how can it be enhanced to bolster resilience?

To answer our question, we set the following three objectives:

1. To describe and visualize the structure of the defense supply chain in the Russian-Ukrainian war.
2. To analyze the roles played by key actors, including manufacturers, NGOs, donors, and conscripts, in sustaining the defense supply chain.
3. To investigate the vulnerabilities of the defense supply chain.

In what follows, we explore civil-military integration across various contexts through a literature review and we review a theoretical framework to describe the levels of the defense supply chain and the roles of key network participants. The method section outlines our approach involving participant interviews and analysis of volunteer funds. In the discussion section vulnerable points in the defense supply chain are identified, and suggestions for strengthening them are offered.

LITERATURE REVIEW

The collaboration between military and civilian entities assumes multifaceted manifestations, each driven by distinct objectives and involving a spectrum of actors, including military personnel, international organizations, local non-governmental organizations (NGOs), and governments (Franke, 2006). We navigate the terminology of Civil-Military Integration (CMI), specifically in the context of the defense supply chain in the Ukrainian case.

NATO's Civil-Military Coordination Doctrine (CIMIC) defines CIMIC as the collaboration between the NATO commander and civil actors in support of missions. Emphasizing cooperation with various organizations, CIMIC's core functions in stability operations encompass liaison, assistance to the civilian environment, and support to the military force (NATO, 2003). Distinguishing itself from Civil Affairs (CA) activities, which doctrinally differ, CIMIC prioritizes humanitarian needs and guides interactions to achieve humanitarian mission objectives, while CA centers on military needs and suggests methods to garner civilian support for military missions (Department of Defense, 2003).

Our study purposefully steers clear of delving into CA in the Russian-Ukrainian war due to the well-established civilian-military relations in Ukraine, which are marked by extensive civilian involvement in supporting and interacting with the Ukrainian military. The Ukrainian military, primarily comprised of civilians mobilized through general conscription, exemplifies CMI in conflict zones. Civilians willingly take on the partial responsibility of supplying resources for the Ukrainian military, thereby enhancing the country's military prowess. This collaborative endeavor is encapsulated by the concept of CMI, aspiring to create a unified and synergistic approach where civilian and military elements work in close tandem, sharing resources, information, and decision-making processes.

CMI emerges as a pivotal facet of collaborative endeavors between civilian and military entities, especially in the realm of global security challenges. Researchers scrutinize CMI from various angles, with one group delving into the theoretical underpinnings and evolutionary dynamics within this context. Brooks (2019) offers an overview

of the theoretical landscape. Su et al. (2020) contributes a quantitative perspective through an evolutionary game framework, and De Coning (2016) delves into the theoretical aspects, shedding light on the complexities and potential constraints in civil-military integration.

A second research perspective regarding CMI revolves around regional insights and case studies, unraveling how this collaboration unfolds in specific contexts. Lafferty et al. (2013) explore China's approach, while Welle (2010) provides an Afghanistan case study, elucidating the challenges and successes in creating a unity of command. Hønneland & Jørgensen (2018) introduce a regional dimension by examining civil-military relations on the Kola Peninsula and the delicate balance between integration and autonomy.

Vuono (2008), Lucius (2016), and Egnell (2009) shift the focus to practical aspects, offering insights into challenges and effective strategies of CMIs. Vuono (2008) addresses practical challenges in stability operations. Lucius & Rietjens (2016) explore practical approaches for cooperation in peace operations, and Egnell (2009) provides insights into navigating diverse operational environments. Berg (2020) investigates the practical implications of civil-military relations in postwar politics, specifically examining their connection to civil war recurrence.

While the existing literature delves into the theoretical underpinnings, challenges, and dynamics of civil-military integration, as well as regional perspectives and case studies, a discernible research gap persists. This gap pertains to understanding the distinct roles, contributions, and challenges of civil initiatives, in contrast to military or government initiatives within the context of CMI. Bridging this gap in the Ukrainian case assumes significance for practitioners seeking effective collaboration strategies, offering a more nuanced understanding of the diverse actors involved and their impact on the success of civil-military integration efforts. Cooperation and possible associated synergies have been modeled with a goal towards quantification in the case of humanitarian organizations by Nagurney and Qiang (2020) using a supply chain network perspective but not yet in a civilian-military defense setting.

THEORETICAL FRAMEWORK

Central to our analysis is the defense supply chain network model that provides a structured framework for understanding the defense supply chain, inspired by Nagurney et al. (2002). The authors introduced a multilevel network perspective for conceptualizing the dynamics underlying supply chains in the presence of competition. This multilevel model comprises the logistical network, the informational network, and the financial network. The authors described the behavior of the decision-makers, including manufacturers, retailers, and consumers.

In contrast, our multi-level structure of the defense supply chain involves four actors: manufacturers, NGOs, and donors. Manufacturers, NGOs, and donors are civilians. NGOs operate voluntarily within this framework. Military conscripts act as the end-users, NGOs serve as intermediaries, manufacturers produce the goods, and donors pay the price of the military product.

We examine the product flow of Unmanned Aerial Vehicles (UAVs) in the defense supply chain – a vital commodity in Ukraine's defensive strategy. The mention of drones serves as an example because even after two years of conflict, drones remain a relevant commodity. Initially, the defense supply chain redistributed goods such as food and clothing, but now there is less demand for such items, whereas drones continue to be in demand. Several studies highlight the significance of UAVs and their applications in military and humanitarian operations. For example, Jasper (2023) discusses the role of UAVs in Ukraine's military response, and Colajanni, Daniele, and Nagurney (2023) construct a three-stage stochastic optimization model focusing on three phases of disaster management integrating 5G technology and UAVs.

Manufacturers

Manufacturers are crucial in determining optimal production quantities and product pricing. Their role is pivotal in shaping supply chain dynamics, as they influence the equilibrium price for UAV products. The manufacturer's supply curve is influenced by the elasticity of production costs. An increase in the supply in UAV production results in higher production costs, affected by factors such as limited inputs (e.g., UAV components and skilled labor).

NGOs

NGOs collect information about the needs of conscripts and inform the civilian community of their needs, purchase goods from manufacturers and deliver them to the military. NGOs determine the volume of shipments based on the conscripts' needs, the manufacturer's price for the product, which they agree upon for the transactions

to take place, and the financial capacity of the community to be donors.

Conscripts

On February 24, 2022, a Presidential Decree declared a general mobilization in Ukraine summoning all military-eligible men, reservists, and conscripts capable of serving based on their health condition to join the armed forces (Presidential Decree, 2022). A part of the Armed Forces of Ukraine consists of men without prior combat experience – conscripts. We examine a scenario in which conscripts were already members of a particular NGO before the onset of Russia's invasion. Positioned as the ultimate consumers in our study, conscripts generate the demand for UAVs based on their military needs. It is important to note that 77% of all NGOs in Ukraine are currently engaged in providing assistance to the Armed Forces (ISAR Ednannia, 2023).

Donors

The process of crowdfunding donations for Ukraine involves leveraging social media platforms such as X (formally Twitter), Facebook, Instagram, and YouTube (Ye et al, 2023). Online channels, facilitated by these platforms, play a crucial role in mobilizing donors for collective action. Donors' financial capacity and relationship proximity with conscripts influence the volume of their contributions, providing an alternative source of income to government funds.

Multi-Level Structure of the Defense Supply Chain Network Model

The defense supply chain model operates on three levels: logistical, informational, and financial (Figure 1).

Logistical Network

- Manufacturers (Top Tier): Producers (m) of UAVs, determining production quantities.
- NGOs (Middle Tier): Intermediaries (n) distributing UAVs to conscripts based on their needs.
- Conscripts (Bottom Tier): End-users (o) with specific affiliations to NGOs, influencing the flow of goods.

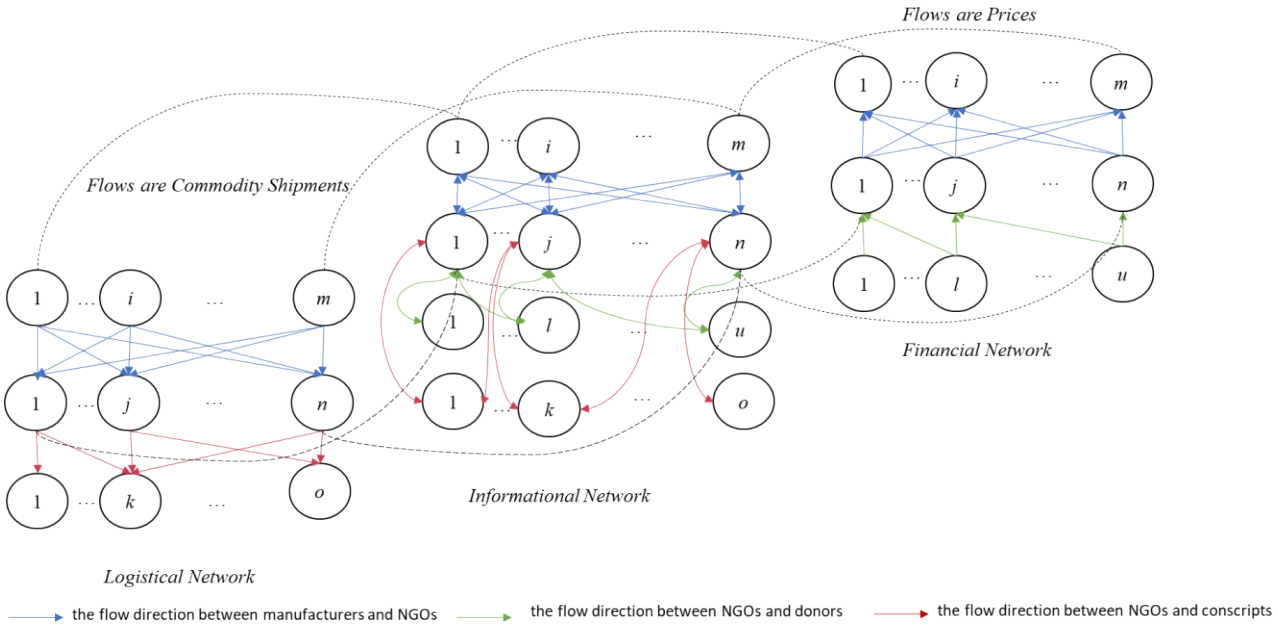


Figure 1. Multi-Level Structure of The Defense Supply Chain Network Model

The logistical network exhibits a downstream flow, starting from manufacturers down to NGOs and ultimately reaching specific conscripts. The flow is contingent on the density of ties between NGOs and conscripts, emphasizing the network's dependence on affiliations.

Informational Network

- Manufacturers (Top Tier): Provide information on prices and quantities of UAVs to NGOs.
- NGOs (Top-Middle Tier): Disseminate conscripts' needs to the civilian community, announce

crowdfunding in local network.
- Donors (Bottom-Middle Tier): Included as an additional level in the information network (u), contribute financial resources based on conscripts' needs announced by NGOs.
- Conscripts (Bottom Tier): Share information about their needs within NGOs' network.

The bidirectional movement of information involves data about commodity shipment quantities and prices, establishing transparency and connectivity within the network.

Financial Network

- Manufacturers (Top Tier): Announce price of UAVs based on costs and availability of goods.
- NGOs (Middle Tier): Facilitate financial transactions by channelling donor contributions to manufacturers.
- Donors (Bottom Tier): Drive purchases by contributing financial resources based on their financial capacity. Donors' income is generated from different sources usually outside the conduct of military operations. The part of income Ukrainian civilians donate regularly for support of an army. A change in income affects the movement of the demand curve.

The upward flow in the financial network reflects a scenario where civilian donors influence purchases through their willingness to contribute, thus driving the financial aspects of the supply chain.

METHOD

To analyze the roles played by key actors and vulnerabilities of the defense supply chain, we use analysed NGO reports and conducted interviews. Quantitative data are obtained from reports issued by the two foremost civilian funds in Ukraine which actively support the Ukrainian military and uphold a commitment to transparent reporting practices – the Serhiy Prytula Foundation, and the "Come Back Alive" Foundation. Concurrently, qualitative insights are garnered through thorough interviews conducted with pivotal stakeholders, encompassing representatives from manufacturing entities, NGOs, and military personnel. These interviews yield nuanced perspectives on decision-making processes, information flow, resource allocation mechanisms, and crucially, the challenges encountered within the defense supply chains.

The methods employed in the research present several limitations that could impact the result. In the context of an ongoing war, the exchange of sensitive information exposes the involved parties to security risks. The interviews were conducted during the summer of 2023. While this timeframe provides valuable insights into the dynamics of the defense supply chain at a specific point in time, it may not reflect changes or developments that occurred before or after the interviews. The ongoing nature of the conflict and the evolving geopolitical situation could influence the effectiveness of the supply chain over time.

The data obtained from reports issued by civilian funds provide valuable information about the flow of goods within the defense supply chain. The "Come Back Alive" Foundation and the Serhiy Prytula Foundation received approximately $286 million USD and $16 million USD respectively in donations between February 2022 and December 2023. However, the scope of these reports may be limited, and they may not capture all relevant aspects of supply chain dynamics.

Despite the limitations outlined in the study, we consider its results valuable. Particularly significant is the portrayal of the ongoing war situation, which can serve as a crucial checkpoint for assessing the broader landscape and dynamics of the conflict. Direct engagement with respondents enabled the establishment of trust and dialogue, enhancing the depth of the findings. Moreover, by using participatory research practices we deem the information valid (Cornwall et al., 1995). One author is in Ukraine, and able to observe the described situation firsthand.

Interviews

To identify connections among actors in the defense supply chain, we conducted interviews. Nine military personnel, three NGO representatives, and one manufacturer were interviewed to gather diverse perspectives on the interactions within the defense network.

The participants included a Ukrainian manufacturer specializing in UAV production. The NGOs selected represented different scales of operation: a nationwide organization, and two local organizations from Lviv and Lutsk (cities in Western Ukraine). These organizations varied in terms of size, annual budget, and the number of members and volunteers in their networks. It is worth noting that all the organizations involved in this study were established prior to 2022, and their primary activities are distinctly different from humanitarian or military operations. Rather, their primary focus is in areas such as education and cultural youth activities (Table 1).

The interviews delved into the methods used by these organizations to disseminate information about the needs of their members, their fundraising strategies, and timelines for collecting and fulfilling those needs.

In addition to inquiring about organizational characteristics and information dissemination strategies, the research methodology encompassed detailed investigations into the types of goods handled by the organizations and the frequency of both receiving and distributing these goods to the military.

Table 1. Characteristics of NGOs Interviewed

	Primary activities	Members	Annual Turnover	Conscripts Supported	Categories of Goods	Source of Funding
Nationwide organization	Education	100-250	Medium (8-40M EUR)	10-50	UAVs, Vehicles, Clothes	International donations, Members' donations
Local NGO (Lviv)	Human Rights	50-100	Small (less 8M EURO)	Less 10	UAVs, Vehicles, Medical supplies	Members' donations
Local NGO (Lutsk)	Youth	less 50	Micro (less 700K EURO)	Less 10	UAVs, Vehicles	Charity Events

Questions addressed organizational membership, categories of goods handled, frequency of transactions, funding sources, storage facilities utilization, duration of processes from need identification to fulfillment, and collaborative partnerships.

The analysis involved synthesizing data from reports with qualitative insights gathered from interviews to understand the interconnectedness and dynamics of the defense supply chain. Findings from both report analyses and interviews were used to draw conclusions and develop recommendations regarding the optimization of the defense supply chain in Ukraine.

FINDINGS

Product Analysis in Defense Supply Chain

In Ukraine, three platforms play a crucial role in supporting the Ukrainian military: the state platform UNITED24, the Serhiy Prytula Foundation, and the "Come Back Alive" Foundation. Notably, UNITED24 serves as a fundraising platform rather than a foundation, channeling contributions directly to relevant ministries, including the Ministry of Defense. The platform has raised $504 million USD in donations from the onset of the conflict until December 2023 (UNITED24, 2023).

The "Come Back Alive" Foundation, operational since 2014, focuses on enhancing the effectiveness of the Ukrainian Defense Forces by procuring equipment such as thermal optics, drones, vehicles, and surveillance systems. The "Come Back Alive" Foundation, received 2,898,936 contributions from February 2022 to December 2023, averaging $100 USD each, with 96% of donations by count and 35% by total amount below $1,000 USD (Come Back Alive, 2023).

The Serhiy Pritula Foundation, established in 2020 in Kyiv, garnered over $16 million USD in donations since the onset of the war, with an average individual contribution of $10 USD. 75% of donations fall within the $1,000 USD range (Serhiy Prytula Foundation, 2023)

Each foundation exhibits unique characteristics in its operations. The "Come Back Alive" Foundation is authorized to allocate funds for weaponry (constituting 20% of its acquisitions) and collaborates with the Ministry of Defense to determine military needs. Additionally, it engages in analytical research that influences nationwide defense decisions and implements projects focused on veteran rehabilitation.

To request assistance for a military unit from the Sergiy Prytula Foundation, an official request must be submitted on the website prytulafoundation.org with the commander's signature and stamp. The foundation then goes through the bureaucratic process, checking for the presence of the stamp and verifying whether individuals are genuinely part of the specified unit. The main challenge arises from the overwhelming number of requests, exceeding the foundation's capacity to fulfill physically (44%). Prioritization becomes necessary based on the unit's location, the tasks it performs, checks on whether aid has been received from other foundations.

Despite variations in reporting practices, our attempt to compare the categories of goods procured by these foundations from January to November 2023 reveals some trends (Figure 2). One notable observation in the foundation reports is the presence of two distinct categories – quadcopters and Unmanned Aerial Vehicles (UAVs), although theoretically, a quadcopter is a type of UAV. Typical quadcopters include the Autel Drone EVO Nano+ and the DJI Mavic 3 Enterprise. UAVs procured by the foundations encompass models such as Leleki, Furii, Valkyrii, among others, each serving diverse military purposes.

Of noteworthy interest, 53% of all acquisitions from the Serhiy Prytula Foundation are comprised of purchases of UAVs and quadcopters (22% and 37% respectively of total acquisitions). In contrast, only 14% of "Come Back Alive" Foundation's acquisitions were UAVs and quadcopters (2% and 12% respectively.

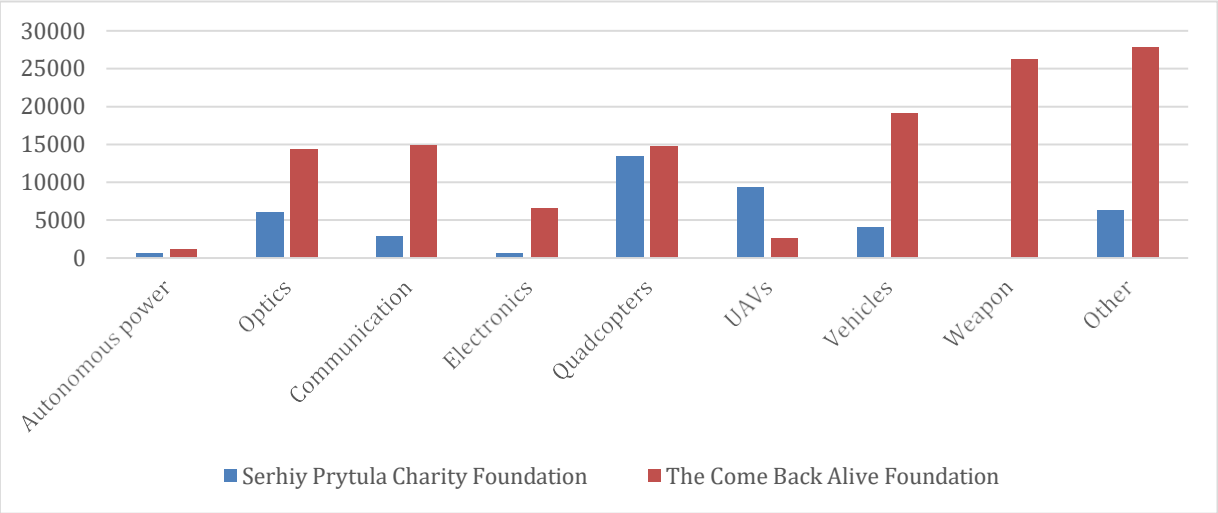


Figure 2. Procurements by Ukrainian foundations during January to November 2023, in thousands of USD

All three organizations interviewed engaged in the collection of funds for UAVs. In addition, one organization also participated in the distribution of medical supplies and another in the distribution of military clothing.

In interviews with military personnel, an additional layer of complexity was uncovered within the defense supply chain. It was revealed that the distribution of needs among military units was not uniform, and this uneven distribution was closely tied to the affiliation of unit members with various civilian networks. The research identified a distinctive pattern wherein the preferences and necessities of military units were intricately linked to their connections within different civilian networks.

An interesting observation emerged: there was a lack of expressed need for food among the military personnel interviewed. Instead, a consistent trend was identified – each individual, regardless of their specific military function, demonstrated a shared experience in expressing the need for fundraising in support of drones and transportation.

Furthermore, military personnel, upon identifying a need, would often use personal savings to fulfill that need before publicly announcing a fundraising campaign within their network of donors. This practice aimed to expedite the satisfaction of requirements by leveraging personal resources before engaging external support networks.

The study also revealed a distinct temporal aspect to the process, with the timeframe for closing requests identified needs ranging from one to three months. This finding shed light on the urgency and efficiency of the defense supply chain in responding to the evolving needs of military units. The dynamic nature of these timelines highlights the adaptability of both the military and the civilian support networks in navigating the challenges posed by the ongoing conflict.

This understanding of the non-uniform distribution of needs, coupled with the intricacies of timing and self-reliance within the military units, contributes valuable insights into the internal dynamics of the defense supply chain and further informs the overarching analysis of the defense supply network. Both report analyses and interviews revealed that UAVs were the most requested item within the defense supply chain; thus, we ground our analyses and discussion around this significant product.

DISCUSSION

Examining the intricacies of the defense supply chain model, we identify key vulnerabilities: network structure vulnerabilities, logistics gap vulnerabilities and financial vulnerability.

Network Structure Vulnerabilities

Non-All-to-All Connections. The defense supply chain operates with a multi-level structure, featuring manufacturers, NGOs, donors, and conscripts. However, the network structure is not an all-to-all connection. This creates challenges in the efficient distribution of information and resources, resulting in a leaky network where gaps in communication and coordination can emerge. Among NGOs, 16,8% continuously involve other NGOs in cooperation after the large-scale invasion and 42,4% in situational project partnerships (ISAR Ednannia, 2023).

Difficulty in Information Flow. The non-all-to-all connection structure makes it challenging to distribute information effectively. This difficulty arises from the fact that a single NGO can assist multiple members, and a recipient can receive aid from various public organizations. The lack of direct connections between all actors hinders seamless information flow, potentially leading to delays and inefficiencies in the supply chain.

Intermediaries as Necessity. To address these challenges, intermediaries become crucial to connect NGOs with one another. Establishing efficient communication and reinforcing networks require that these intermediaries facilitate the exchange of information and resources. Identifying and leveraging influential actors becomes essential for enhancing connectivity within the network.

Logistics Gap Vulnerabilities

Under-Responsiveness. The defense supply chain is vulnerable to under-responsiveness, especially in scenarios where the logistics network fails to provide the right mix of resources at the right place and time. This under-responsiveness results in a gap between the available and required logistic resources, impacting the effectiveness of combat operations.

Cumulative Shortfalls. Shortfalls in resource availability accumulate over time, leading to a widening logistics gap. If under-responsiveness persists, this gap can reach a critical point known as the logistic culmination point. Beyond this point, combat operations cannot be executed as planned, and the effectiveness of military endeavors decreases significantly.

Financial Vulnerability

Funding is essential to NGOs engaged in humanitarian operations and, for many such organizations, securing financial donations is critical to their very sustainability (cf. Toyasaki and Wakolbinger (2014), Nagurney et al. (2018), Nagurney, Salarpour, and Daniele (2019)).

The financial vulnerability of the supply chain is intricately linked to the absolute inelastic demand for military goods, particularly for UAVs. With a price elasticity of zero, changes in the product price do not influence the quantity demanded. The demand function, dependent on conscript needs and donors' income, shapes the dynamics. Conscript needs, influenced by attrition and battlefield consumption, are crucial determinants.

The financial vulnerability arises when the demand surpasses the available supply, triggering an increase in price and production volume. However, internal constraints also exist, demanding a simultaneous growth in donor incomes. If donor incomes stagnate while needs rise, a temporary surge in demand occurs, escalating prices and reducing donor incomes even further. Similarly, if needs outpace declining donor incomes, sustained demand necessitates even higher donor income growth, highlighting the delicate financial equilibrium within the civilian-military supply chain.

Resilience in the Defense Supply Chain

The vulnerabilities in the defense supply chain are rooted in the network structure's complexity, the financial vulnerability, and the potential for under-responsiveness, leading to a logistics gap. Addressing these vulnerabilities requires a focus on strengthening connections, leveraging intermediaries, supporting income growth for donors, and ensuring the adaptability of the logistics system to maintain a responsive and resilient supply chain (Figure 3).

Communication line for NGOs. Efficient communication and network reinforcement require intermediaries to connect NGOs with one another. Identifying influential actors helps to pinpoint organizations that can serve as

effective mediators in the network.

Donor Income Growth. In the realm of financial vulnerability, resilience hinges on the crucial alignment of donor income growth with the escalating prices for resources and demand. The financial network highlights the pivotal connection between donor income and the demand functions, emphasizing the need for a symbiotic relationship where increased demand and prices stimulate growth in donor incomes.

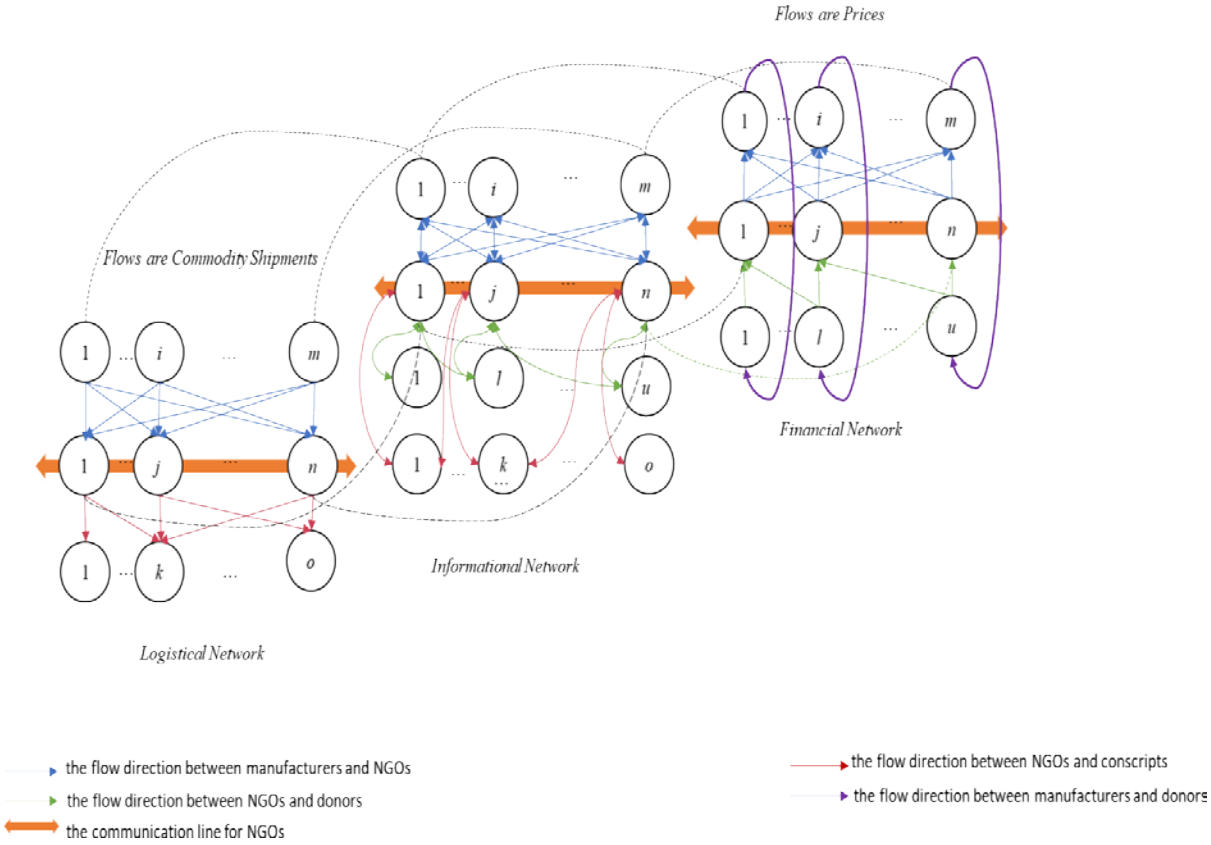


Figure 3. Resilient Defense Supply Chain Model

A noteworthy facet of resiliency involves donors working with the same manufacturers that produce the needed goods, such as UAVs. Procurements by Ukrainian foundations during the war are noteworthy for their sources of origin. For instance, UAVs Leleki, Furii, and Valkyrii are manufactured by Ukrainian companies, while quadcopters Autel and Mavic are of Chinese origin. It is important to mention that in Ukraine, approximately 200 drone manufacturers produce various types of drones, but, these manufacturers often rely on components from Chinese production (Balashova & Melnyk, 2023). In this dynamic, when donors obtain income from external sources, they contribute an amount equivalent to what is accessible as donations. However, as prices rise, their donations decrease, creating a delicate equilibrium between sustaining demand, prices, and donor incomes. The resilient defense supply chain underscores the imperative of directing financial resources towards Ukraine's military-industrial complex. This strategic allocation ensures stability through consistent salaries for the Ukrainian population.

CONCLUSION

We set out to answer the question: To what extent is the defense supply system in Ukraine vulnerable, and how can it be enhanced to bolster resilience? To answer this question, we used three objectives. Beginning with Objective 1 we described and visualized the structure of the defense supply chain in the Russian-Ukrainian war (Figure 3). The multilevel defense supply chain developed provided insight regarding key actors and illuminated their interactions within the supply chain, spanning logistical, informational, and financial dimensions (Objective 2). Our analysis revealed the pivotal roles played by manufacturers, NGOs, donors, and conscripts in sustaining the defense supply chain. By employing both quantitative data from reputable civilian funds and qualitative insights from interviews with key stakeholders, we gained an understanding of the dynamics at play. Finally in our third objective, we sought to investigate the vulnerabilities of a defines supply chain. Our investigation

identified several vulnerabilities including network structure complexities, logistics gaps, and financial fragilities. These vulnerabilities underscore the need for strategic interventions aimed at strengthening connections, leveraging civilian intermediaries, and ensuring the adaptability of the logistics system.

Importantly, our study emphasizes the importance of civilian-military integration in bolstering a nation's military capabilities. By recognizing civilians as active contributors to the defense supply chain, rather than mere recipients of aid, we advocate for a more unified approach that maximizes resources and enhances overall effectiveness.

Moving forward, addressing the identified vulnerabilities will be crucial for enhancing Ukraine's defense resilience. By channeling financial resources into the military-industrial complex and promoting collaboration between civilian and military entities, Ukraine can build a more robust defense supply system capable of effectively responding to evolving challenges and safeguarding its sovereignty.

As further work, we aim to expand our focus to encompass the development of an information network for interaction initiatives within the defense supply chain, represented by NGOs. By examining the interdependence between the flow of goods and the density of connections among NGOs, we seek to identify opportunities for enhancing efficiency, mitigating risks, and improving coordination within the defense supply chain. By undertaking these additional research efforts, we aim to deepen our understanding and identify practical measures to address vulnerabilities and enhance the resilience of the supply chain.

REFERENCES

Balashova, L., & Mel′nyk, T. (2023). Руки для крил. Українські компанії виготовляють близько 50 000 FPV-дронів щомісяця, потрібно – сотні тисяч. Як кадровий голод гальмує галузь (in Ukrainian). *Forbes* Available at:. https://forbes.ua/innovations/ukrainski-inzheneri-pratsyuyut-u-boeing-ta-airbus-virobniki-droniv-strazhdayut-cherez-defitsit-kadriv-yak-rozvyazati-tsyu-problemu-20112023-17379. Last Accessed: 30/12/2023.

Barany, Z. (2023). Armies and autocrats: Why Putin's military failed. *Journal of Democracy*, 34(1), 80-94. DOI:10.1353/jod.2023.0005.

Berg, L. A. (2020). Civil-military relations and civil war recurrence: Security forces in postwar politics. *Journal of Conflict Resolution*, 64(7-8), 1307-1334. DOI: 10.1177/0022002720903356

Brooks, R. A. (2019). Integrating the civil–military relations subfield. *Annual Review of Political Science*, 22, 379-398. DOI:10.1146/annurev-polisci-060518-025407

Channel-Justice, E., (2022). *Without the State: Self-Organization and Political Activism in Ukraine*. University of Toronto Press, Toronto, Canada.

Colajanni, G., Daniele, P., & Nagurney, A. (2023). Centralized supply chain network optimization with UAV-based last mile deliveries. *Transportation Research C*, 104316.

Colajanni, G., Daniele, P., Nagurney, A., Nagurney, L.S., & Sciacca, D. (2023). A three-stage stochastic optimization model integrating 5G technology and UAVs for disaster management. *Journal of Global Optimization*, 86, 741-780.

Come Back Alive (2023, December 1). Fund Reporting. Available at: https://savelife.in.ua/en/reporting-en/. Last Accessed: 30/12/2023.

Cornwall, A., & Jewkes, R. (1995). What is participatory research? Social Science & Medicine, 41(12),1667–1676. https://doi.org/10.1016/0277-9536(95)00127-S

Cullen Dunn, E., & Kaliszewska, I. (2023). Distributed humanitarianism: Volunteerism and aid to refugees during the Russian invasion of Ukraine. *American Ethnologist*, 50(1), 19-29.

Dalsjö, R., Jonsson, M., & Norberg, J. (2023). A brutal examination: Russian military capability in light of the Ukraine war. *Survival*, 64(3), 7–28.

De Coning, C. (2016). Civil-military interaction: Rationale, possibilities and limitations. In: Lucius, G., Rietjens, S. (eds) *Effective Civil-Military Interaction in Peace Operations*. Springer, Cham.

Department of Defense (2003). “Joint Doctrine for Civil Affairs—JP 3-57.1”. Available at: http://www.dtic.mil/doctrine/jel /new_pubs/jp3_57_1.pdf. Last Accessed: 30/12/2023.

Egnell, R. (2009). *Complex Peace Operations and Civil-military Relations: Winning the Peace*. Routledge.

Franke, V. (2006). The peacebuilding dilemma: Civil-military cooperation in stability operations. *International Journal of Peace Studies*, 5-25.

Hackett, J. (2023). The shadow of war. The International Institute for Strategic Studies. Available at: https://www.iiss.org/online-analysis. Last Accessed: 30/12/2023.

Hoffman, F. (2022). Defining and achieving success in Ukraine. *PRISM*, 10(1), 88-105.

Hønneland, G., & Jørgensen, A. K. (2018). *Integration vs. Autonomy: Civil-military Relations on the Kola Peninsula*. Routledge.

International Institute for Strategic Studies. (2021). *The Military Balance 2021*. Routledge.

ISAR Ednannia (2023). Ukrainian civil society under the war – 2022. Retrieved December 30, 2023, from https://ednannia.ua/attachments/article/12447/Ukrainian%20civil%20society%20under%20the%20war.pdf. Last Accessed: 30/12/2023.

Jasper, S. (2023). Resilience against hybrid threats: Empowered by emerging technologies: A study based on Russian invasion of Ukraine. In Balomenos, K., Fytopoulos, A., & Pardalos, P. M. (eds) *Handbook for Management of Threats: Security and Defense, Resilience and Optimal Strategies*. Cham: Springer International Publishing.

Konrad, R., Sorokotyaha, S., & Walker, D. (2023). Humanitarian response by grassroots associations during a military conflict. *Journal of Humanitarian Logistics and Supply Chain Management*. In press.

Lafferty, B., Shraberg, A., & Clemens, M. (2013). China's civil-military integration. *SITC Research Briefs*, 2013(2013-10).

Ley, S. (2022). High-risk participation: Demanding peace and justice amid criminal violence. *Journal of Peace Research*, 59(6), 794-809.

Lucius, G., & Rietjens, S. (2016). *Effective Civil-Military Interaction in Peace Operations*. Switzerland: Springer International Publishing.

Melzer, N., & Kuster, E. (2019). International humanitarian law. A comprehensive introduction. Geneva: International Committee of the Red Cross.

Nagurney, A. (2023). Defense-critical supply chain networks and risk management with the inclusion of labor: Dynamics and quantification of performance and the ranking of nodes and links. In Balomenos, K., Fytopoulos, A. & Pardalos, P. M. (eds) *Handbook for Management of Threats: Security and Defense, Resilience and Optimal Strategies*. Cham: Springer International Publishing.

Nagurney, A., Daniele, P., Alvarez Flores, E., & Caruso, V. (2018). A variational equilibrium network framework for humanitarian organizations in disaster relief: Effective product delivery under competition for financial funds. In Kotsireas, I., Nagurney, A. & Pardalos, P.M. (eds) *Dynamics of Disasters: Algorithmic Approaches and Applications*. Springer International Publishing Switzerland.

Nagurney, A., Ke, K., Cruz, J., Hancock, K., & Southworth, F. (2002). Dynamics of supply chains: A multilevel (logistical–informational–financial) network perspective. *Environment and Planning B: Planning and Design*, 29(6), 795-818.

Nagurney, A., & Qiang, Q. (2020). Quantifying supply chain network synergy for humanitarian organizations. *IBM Journal of Research and Development*, 64(1/2), 12:1-12:16.

Nagurney, A., Salarpour, M., & Daniele, P. (2019). An integrated financial and logistical game theory model for humanitarian organizations with purchasing costs, multiple freight service providers, and budget, capacity, and demand constraints. *International Journal of Production Economics*, 212, 212-226.

NATO (2003). NATO Civil-Military Co-Operation (CIMIC) Doctrine AJP-9. Available at: www.nato.int/ims/ docu/AJP-9.pdf. Last Accessed: 30/12/2023.

Presidential Decree No 65/2022 "On General Mobilization". Available at: https://www.president.gov.ua/ documents/652022-41653. Last Accessed: 30/12/2023.

Serhiy Prytula Foundation (2023, December 1). Fund reporting. Available at: https://prytulafoundation.org/en/military-reports. Last Accessed: 30/12/2023.

Slim, H. (2022). Humanitarian resistance: Its ethical and operational importance. *Humanitarian Practice Network*, 87.

Su, Z., Yang, X., & Zhang, L. (2020). Evolutionary game of the civil-military integration with financial support. *IEEE Access*, 8, 89510-89519.

Ti, R., & Kinsey, C. (2023). Lessons from the Russo-Ukrainian conflict: The primacy of logistics over strategy. *Defence Studies*, 23(3), 381-398.

Tierney, D. (2023). How Putin tore up his playbook for victory. *The Washington Quarterly*, 46(3), 185-

200. Toyasaki, F., & Wakolbinger, T. (2023). Impacts of earmarked private donations for disaster

fundraising.
Annals of Operations Research, 221(1), 427-447.

UNITED24 (2023, December 1). Fund reporting. Available at: https://u24.gov.ua/uk. Last Accessed: 30/12/2023.

Vuono, T. A. (2008). Challenges for civil-military integration during stability operations. US Army War College - USAWC civilian research paper.

Welle, J. W. (2010). Civil-military integration in Afghanistan: Creating unity of command. *Joint Force Quarterly*, 1(56), 54-59.

Ye, J., Jindal, N., Pierri, F., & Luceri, L. (2023). Online networks of support in distressed environments: Solidarity and mobilization during the Russian invasion of Ukraine. *arXiv preprint arXiv:*2304.04327. Last Accessed: 30/12/2023.



Usability and Universal Design of ICT for Emergency Management

Track Chairs

Terje Gjøsæter, University of Agder

Jaziar Radianti, University of Agder

Laura Petersen, UIC Paris

Weiqin Chen, Oslo Metropolitan University

Contextualizing Effective Use of Health Information Systems in Humanitarian Setting: The role of Transparent Interaction

Marta María Vila-Pozo
HISP Centre, University of Oslo
Médecins Sans Frontières Spain
martamv@ifi.uio.no

ABSTRACT

Humanitarian organizations strive for efficient and effective interventions, with information systems playing a crucial role. However, integrating data collection and processes within their operations remains a challenge. This paper examines the effective use of a Health Management Information System in a humanitarian context. Effective use aims at helping information systems contribute to the organisational objectives and holds promises to help improving information systems in this domain. The analysis uncovers obstacles related to system access and learning processes, suggesting that transparent interaction is vital for success. Transparent interaction is the first dimension in the Theory of Effective Use. Through contextualization, this paper defines effective use in humanitarian settings and evaluates the theoretical model's validity in this unique context. Drawing from concepts from the information systems body of theory, it contributes to the understanding of effective use by offering a contextualized model and a clear definition of transparent interaction, emphasizing its significance.

Keywords

Effective Use, Humanitarian Health Information Systems, Transparent Interaction, Affordance Network.

INTRODUCTION

The past year has been marked by ongoing conflicts, severe climate issues, and health crises like cholera and COVID-19. These factors have led to an alarming rise in hunger, displacement, and poverty. As a result, it is estimated that 339 million people will require humanitarian assistance and protection. This is a significant increase from the 274 million people estimated at the start of 2022. In short, one out of every 23 people will be in need of humanitarian aid (OCHA, 2022). In response to such existing and increasing demand humanitarian organizations assist people and communities who are in need and face significant challenges. This includes refugees, asylum - seekers, internally displaced populations, and others who have been affected by diseases, conflicts, and both human-made and natural disasters. Humanitarian organizations need to ensure efficiency and efficacy in their interventions and are expected to share data effectively to provide more coordinated interventions, imp rove accountability and provide efficient response (Checchi et al., 2017). The use of digital technologies and information systems are an integral part of the deployment of the response provided (Hunt et al., 2016). They facilitate communication, coordination, and the collection and analysis of large amounts of data, enabling timely responses in humanitarian contexts (Mesmar et al., 2016).

Information systems are contributing to providing the required information, and operational improvements involve the use of project management information systems and digital technologies for data collection and analysis. This shift has been observed in many humanitarian organizations over the past decade (ALNAP, 2015), replacing outdated spreadsheets and paper-based systems. These changes are expected to increased efficiency and effectiveness in their interventions; however, organizations have not yet fully succeeded in integrating data

collection and integration processes within their operations (ALNAP, 2022). Field interventions are not heavily scrutinized form a public health perspective among humanitarian actors, however, public health data is fundamental for documenting processes and interventions and to reinforce accountability and advocacy, really important in the context (Jarrett et al., 2021). A recent study of the status of the humanitarian system on the evidence of monitoring and evaluation data from eight big humanitarian funders showed evidence of impact was only available for 16% of the projects (ALNAP, 2022), in most cases it was not possible to collect data on the outcomes after the funding ended. This reveals a significant data gap on what impact innovations are collectively having on humanitarian effectiveness and efficiency (ALNAP, 2022).

The role of Humanitarian Health Management Information Systems (H-HMIS) is to collect routine data from the field interventions and generate the indicators and data aggregations necessary to follow up the status of the intervention and population needs (Magnuson & Dixon, 2020). Routine health data provides crucial public health information, enabling real-time surveillance, monitoring systems, and guiding timely interventions. Additionally, it facilitates the evaluation of intervention and accountability towards the organization's mission and objectives (Checchi et al., 2017). Measuring impact in the wellbeing of the population might exceed the domain of information systems theories, however, having robust information systems is a key component for it. The theory for Effective Use holds promises to contribute to improving information systems in this domain by helping assess to what extent information systems are effectively contributing to the organisational objectives and provide instrumental guidance to organizations to strengthen their internal information systems and report ing mechanisms. With the aim of analysing the adequacy of the theory for the context of humanitarian organizations, this paper aims to answer the following research question:

What are the challenges for attaining effective use of a humanitarian health management information system?

In the empirical setting of a field intervention of Médecins Sans Frontières Spain (MSF Spain), this analysis builds upon previous research which has studied the implementation and adoption of a H -HMIS from the headquarters offices in the global context of the organization (Vila-Pozo & Sahay, 2018), to the project coordination office (Vila-Pozo & Sahay, 2019) and the inpatient wards in the field hospitals (Vila-Pozo et al., 2022) of one of its field interventions.

The core analysis is based on a contextualization of the theory for effective use. In the next section, the theoretical component of this analysis is introduced. The third section introduces the case study, followed by an analysis and findings. The paper concludes with the discussion of findings and theoretical contributions.

RELEVANT LITERATURE

Theory of Effective Use

The theory of effective use (TEU) draws from Representation Theory (RT) and the conception of information systems as semantic systems (Wand & Weber, 1995). Representation Theory conceptualizes systems as capable of representing phenomena from which individuals retrieve digital representations of the real world. Therefore, an information system conveys meaning through a structure of signs. Representation Theory posits that information systems are comprised of three interrelated structures:

- Deep structure: This refers to the definition of data structures and their specifications to represent the domain. In the case of the H-HMIS it refers to the database and data model definition.
- Surface structure: This is the structure that enables the user to access the deep structure. It would mainly be the user interface of the H-HMIS.
- Physical structure: This component connects the deep and surface structures. This is the hardware supporting the implementation, from the computers or mobile devices to the broader infrastructure.

Individuals use digital representations to gain knowledge about a particular situation from the real world that informs their actions (Burton-Jones et al., 2017). The transition from use to effective use can be seen as the evolution from looking at the mere use of the system to perform a specific task (system use) to its use in a manner that contributes to the attainment of the organizational desired goals (effective use). The use of the information system is not seen as a goal itself, but as a mean to an end, i.e., the organization's objectives. Effective use is defined as "using a system in a way that helps attain the goals for using the system " (Burton-Jones & Grange, 2013). The nature of effective use is modelled across the following three dimensions, as described by Burton - Jones and Granje (2013):

- Transparent interaction: refers to "the extent to which a user is accessing the system's representations unimpeded by its surface and physical structures".

- Representational fidelity: relates to "the extent to which a user is obtaining representations from the system that faithfully reflect the domain being represented".
- Informed action: refers to "the extent to which a user acts upon the faithful representations he or she obtains from the system to improve his or her state in the domain of work".

The relationship between these dimensions is hierarchical. As a hierarchical network, transparent interaction activates information potential, which enables representational fidelity. Representational fidelity can then be leveraged by informed action (Burton-Jones & Grange, 2013). Actualizing these three dimensions is crucial to achieve effective use, where each of them is necessary but not sufficient for the next. This relationship is visually represented in Figure 1 below.

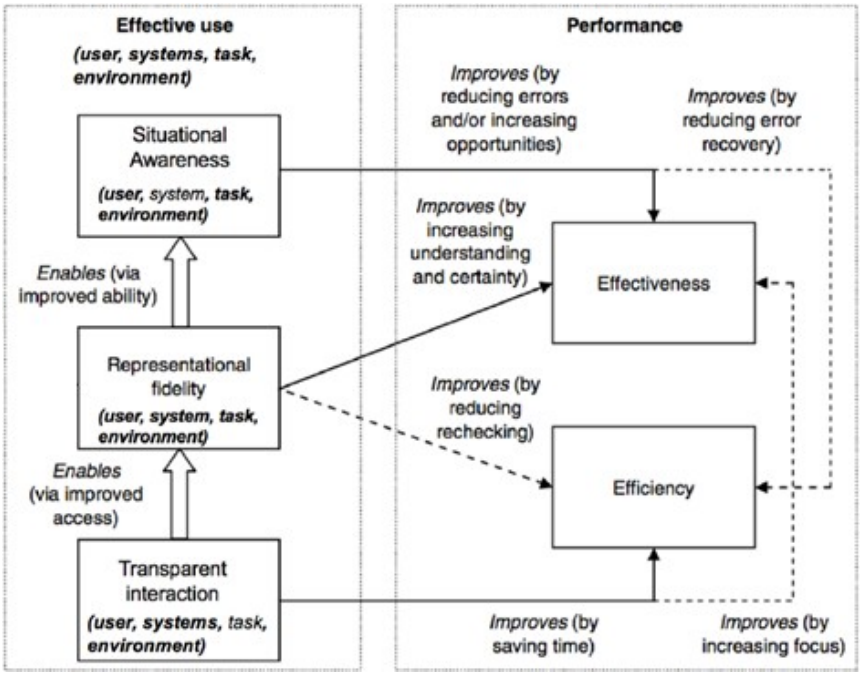


Figure 1. Effective Use model adapted for emergency management contexts (Bonaretti & Piccoli, 2018).

The theory of effective use has been used to guide research on IS implementations in a variety of domains, such as hospitals (Eden et al., 2018; Eden et al., 2019), educational providers (Eden et al., 2020), enterprises (Trieu, 2013), banking (Haake et al., 2018), and more recently in Emergency Management (Bonaretti & Piccoli, 2018) which proposed modifications to the TEU model. The main modification involved replacing the Informed Action dimension with Situational Awareness as the third dimension of effective use. Situational awareness assumes that all the necessary information may not be present in the information system (Endsley, 1995), and the focus instead is on sufficiency of the information as a foundation of effective use. The resulting TEU model adapted for emergency management contexts can be also found in Figure 1. When analysing the dimensions for effective use in this paper, I will use this model with situational awareness as the third dimension and the environment included as a function for effective use in addition to the user, the system, and the task, that are introduced in the original model. My application of the theory of effective use is based on its operationalization for the humanitarian field .

CASE ANALYSIS AND FINDINGS

Research Context

This work builds upon a longitudinal case study (2015-2021) examining the implementation and adoption of the H-HMIS across the whole organization. From the headquarters offices in the global context of the organization (Vila-Pozo & Sahay, 2018), to the project coordination offices (Vila-Pozo & Sahay, 2019) and further to inpatient wards in two field hospitals (Vila-Pozo et al., 2022) within one of its field projects deployed to providing health care in a camp for internal displaced populations.

The empirical setting of this analysis was the field mission of MSF Spain in South Sudan. South Sudan has been a consistent focus for MSF Spain since 2013, following the eruption of civil war shortly after the nation's 2011

independence. In the ensuing chaos, an estimated 400,000 lives were lost, 2.28 million individuals sought refuge in neighbouring countries, and another 1.87 million were displaced within South Sudan's borders (MSF Spain, 2021).

Analysis and Findings

This section presents a contextual analysis based on the conceptualization of the theory for effective use in a humanitarian setting. The approach follows a step-by-step recommendation from Burton-Jones and Volkoff (Burton-Jones & Volkoff, 2017) for developing contextualized theories of effective use which consists on:

1. Understanding how a network of affordances supports the achievement of organizational goals.
2. Understanding how the affordances are actualized.
3. Using inductive theorizing to elaborate those principles in a given context.

The theory of effective use, and this process for its contextualisation, govern the analytical framework guiding this research. This approach however, introduces and builds upon the concept of affordances (Strong et al., 2014) and network of affordances (Burton-Jones & Volkoff, 2017). Affordances are theoretically defined as "the potential for behaviours associated with achieving an immediate concrete outcome" (Strong et al., 2014, p. 69), and their actualization is the enactment of such potential in context. Affordances, hence, refer to the relation between the users and the system and their capacity to use it to support their everyday actions, in our case related to the use of the information system during a humanitarian intervention. In the interest of clarity and to align with the theoretical emphasis of this article on the theory of effective use, the term affordances will, in the context of this study, be substituted with *functionalities* and their related process of actualization or realization. Consequently, the network of affordances, will be designated as the network of functionalities facilitating the attainment of the organizational objectives. The following sections develop each of the three different phases enumerated above, using the proposed terminology.

Understanding how a Network of Functionalities Supports the Achievement of Organizational Goals.

The network of functionalities is built upon outcome-units, which represent the relationship between each functionality and its potential outcome (Burton-Jones & Volkoff, 2017). These outcome-units collectively contribute to the overall organizational objectives. By analysing the interconnections between functionalities, the network allows us to understand the interdependence, chronological sequence, and presence of functionalities and outcomes across different organizational levels. Taking together the functionalities and organizational objectives, the analysis helps to understand how the technology helps to meet the objectives of the organization, concerning the effective use of the system in the field project. As introduced above, prior research has identified the operating functionalities of an H-HMIS at the two field levels under study *i)* the project coordination office (Vila-Pozo & Sahay, 2019), and *ii)* the inpatient ward in the health facilities (Vila-Pozo et al., 2022).

Table 1 Operating Functionalities of H-HMIS in the field setting

Level	Functionality	Description
Project Coordination	Operationability	The possibility to use data to follow up trends and identify alerts for constant response and adaptation of activities.
	Contextuability	the possibility to know and analyse context when you are new to the intervention.
	Accountability	the possibility of data to be accountable for day-to-day activities and operational decisions.
Health Facility [1]	Collectability	The possibility to collect patient data along with the medical interventions.
	Aggregability	The possibility to aggregate data at the end of each reporting period.
	Manageability	The possibility to manage the hospital ward with the data collected.
	Recordability	The possibility to access patient historical records when they are readmitted.

[1] The functionalities at facility level were published with a different naming convention. Names have been adjusted to match the style of those identified at the project coordination level.

The organizational objectives to which the functionalities contribute to are those of:

Monitoring the quality of health services: this objective refers to the continued follow-up, assessment, and evaluation of the different aspects of the intervention related to the health services provided. The functionalities contributing to this objective are mainly those operating at the project coordination office. Operationability allows the tailoring of the services provided to the actual and changing needs. Accountability ensures following the organizational paths for mobilizing or relocating resources, and contextuability helps newcomers gain understanding of the situation when initiating their professional activity in context.

Minimizing data management workload of field workers: this objective aims at reducing the effort and time dedicated by field health staff to collect and process data. The functionality contributing principally and directly to this objective is that of Aggregability, since it eliminates the manual processes of aggregation and subsequent data entry, considered the most time consuming and prone to errors by field staff (Vila-Pozo & Sahay, 2019).

Reducing reporting and data integration delays: this objective aims at reducing the time span between data collection and its availability for integrated analysis at the higher levels in the organization. The functionality described as aggregability is again the main one contributing to this objective as it eliminates both, the manual process of aggregation and the data entry of the routine information into the H -HMIS. Once the data reaches the system, its integrated architecture guarantees data available at all levels at close to real-time without manual steps at the higher levels. Being data aggregation, data entry and data integration the main cause of delays in having managerial information, aggregability becomes the key functionality contributing to this objective.

Improving data use and informed decision-making: this objective aims at providing access to data and information and the right time to support operational decisions. From a coordination level, all three functionalities build the basis for informed decision-making for managing the intervention. At a lower level, manageability contributes to decision making at a lower level for the day-to-day logistics of the ward.

The network of functionalities

The representation of the network of functionalities presented below helps to illustrate the dependencies between functionalities and their contribution to achieving organizational goals. The diagram uses different colours to distinguish between facility-level (green) and project coordination-level (blue). The arrows between levels indicate the dependencies and follow the colour of the source level.

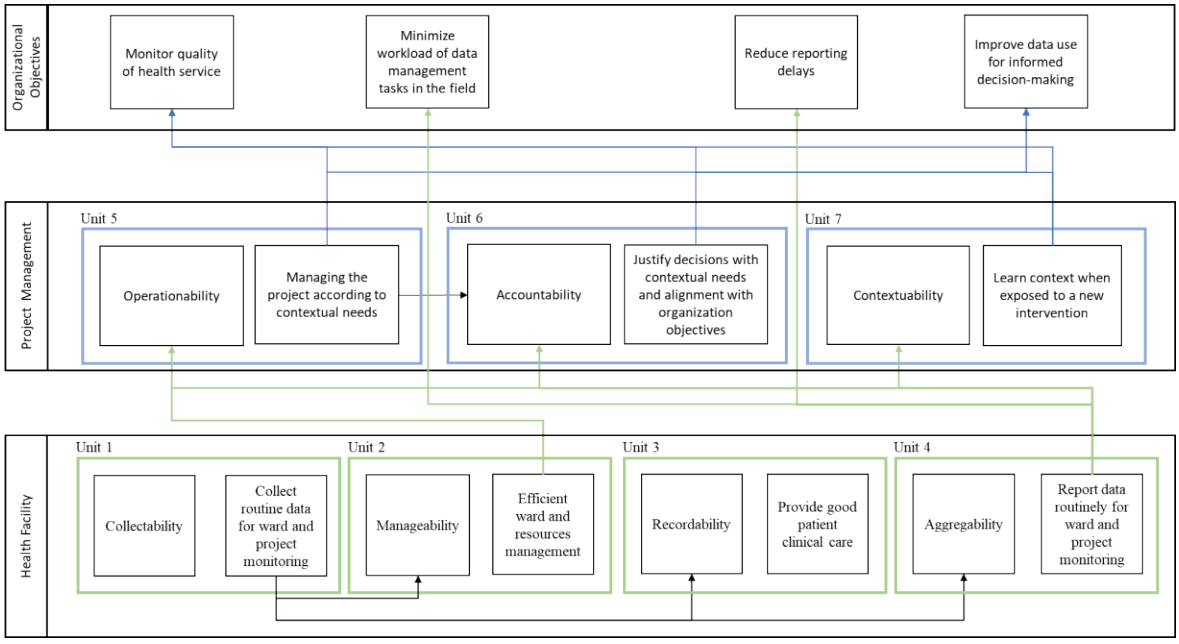


Figure 2. The network of functionalities.

This analysis focuses on the data collection process in a H-HMIS, from four different in-patient wards: adults, paediatrics, neonatology, and therapeutic feeding program, to the coordination office. In all these wards, patient data is collected daily through registration in a register book, where each patient is registered in one line with the basic information (Unit 1). This data serves as a basis for managing the ward (Unit 2), including the provision of food supplies, pharmacy stocks, bed linen, and bed nets, as well as accessing patient information for future readmissions (Unit 3). Additionally, at the end of each reporting period, data is aggregated on a weekly basis to

generate project monitoring reports (Unit 4). These outcome-units operate at the health facility level, they enable the functionalities at the next level and, in some cases (Unit 2 and 4), contribute directly to the organisational objectives of minimizing workload of data management in the field and reduce reporting times.

At the project coordination level, operationability is crucial to manage the evolution of the project according to contextual needs (Unit 5) and enables the functionality described as accountability (Unit 6). Contextuability allows new staff members to familiarize themselves with the specific work environment (Unit 7). The three functionalities contribute to the continued monitoring of quality of the health service and the improvement of data use for enabling informed decision-making.

Understanding how the Functionalities are Actualized

Building upon the network of functionalities as the path for effective use, the next analytical step is the study of the process of actualization for this setting and its work conditions.

To determine the potential impact of the context on achieving effective use, the next step involves analysing the operating conditions. The contextual characteristics operating in the deployment of an H-HMIS in a humanitarian setting respond to those of *dynamicity*, *restrictive infrastructure* that requires *independency*, and *heterogeneity* (Vila-Pozo & Sahay, 2018). Those contradict with the fundamental principles governing routine health information system: stability, infrastructural dependency, and homogeneity (Vila-Pozo & Sahay, 2018).

To better understand the impact of such contradictory principles on achieving effective use, it is necessary to analyse the operating contradictions at a lower level, specifically with respect to the core components of the information system: the deep, surface, and physical structure (Wand & Weber, 1995).

Deep structure: The H-HMIS evolves in a context where work routines tend to not facilitate the requirements gathering process, which potentially results in inadequate database and design configurations. In addition, the varied modes of operation through which humanitarian organizations must respond to different emergencies, such as armed conflicts, natural and man-made disasters, and epidemics raise different information needs. This requires a flexible data model which directly shapes the deep structure of the information system.

Surface structure: The context in which the system is implemented is also dynamic and the needs of the population evolve as the emergency unfolds. This creates a need for different data to be collected, and possibly from different new locations. The surface structure, consisting of the user interface, must be adaptable and able to absorb these changes to provide a meaningful user experience that facilitates data collection in multiple settings. In addition, there are severe infrastructure limitations that raise the need to support offline mode, data synchronization, and distributed systems, to manage the distributed data flows. The need to operate in independent and restricted settings can potentially affect the surface structures of the information system that needs to offer a flexible yet simple user experience.

Physical structure: The dynamic context explained above creates a need for different data to be collected from new locations. The physical structure may be affected due to the opening (and closing) of health units, with varying levels of access to Internet or electricity networks. This combined with environments with severe infrastructure restrictions where internet access, regular electricity supply, and where adequate workstations may not be readily available poses challenges on hardware and infrastructu re. Deploying an information system that can function under such conditions directly affects the physical structure. The interplay between the system principles, the operating conditions and their impact on the system structures is represented in the figure below.

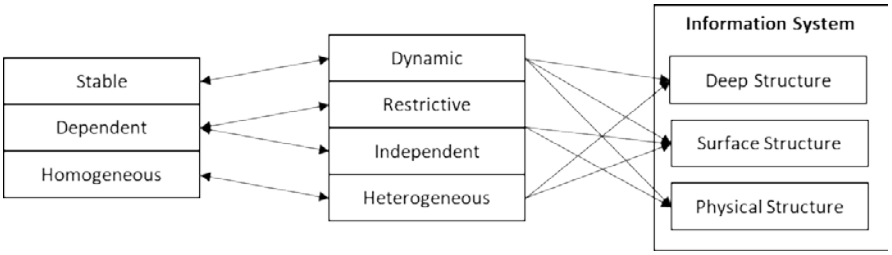


Figure 3. Conflicting logics and influence on system structures.

The environment, hence, affects all structures of the health information system, which in turn influence its implementation adoption and use. The operating conditions shape both, the features of the system as well as the interaction system-user. These characteristics of the humanitarian environment generate what I conceptualised as operating conditions. The operating conditions shape the actualization of the functionalities at their different levels of operation, sometimes as enablers sometimes as blockers for system use.

They have been identified as (Vila-Pozo & Sahay, 2019):

Mature use of information and positive perception of data was shown on day-to-day work practices. The field team relied on data to inform their meetings and decisions. Team leads encouraged collaborative analysis and events were discussed basing their arguments in the data available, when possible.

Limited available information refers to the cases where the configuration of the H-HMIS in terms of the health content available to collect and analyse was not flexible enough to cover needs of the field teams leaving out sometimes critical information, while including other aspects less relevant for the intervention.

Characteristics of the humanitarian work environment represented the biggest challenges. The work environment is intense, and the team works long hours- Staff is limited, and personal and professional spaces and times are often mixed. Turnover rates among field staff are high which adds complexity to the already challenging work atmosphere.

Technical features of the system at the infrastructure level, require having an offline server which operates as independently as possible from the internet connection and power supplies, which influences both the infrastructure and hardware as well as the user interface. At a functional level, the automatic integration of data in one central database was seen as useful for integrated analysis. The system, however, was perceived as adequate for data entry tasks, but having all data in one system was perceived as complicated by users accustomed to work on excel files for data analysis.

The impact of such operating conditions shapes the actualization of functionalities when the user interacts with the technology. At the project coordination office, the actualization process suffered interaction effects – enabling or constraining, which shaped the overall actualization of the functionalities (Vila-Pozo & Sahay, 2019). The limitations in information and the work environment where always constraining and negatively affecting the actualization of functionalities, as they imposed challenges at different levels, from data related limitations to human resources difficulties caused by a high turnover of staff, very common in these types of organizations.

Other operationalconditions had dual effects. The mature use of data was perceived as fuel for system use is most cases, however, when combined with limitations in the information available, resulted in a motivation for using parallel tools and ad-hoc solutions, including does based on paper reporting. The technical features had similarly dual effects, a centralized integrated model introduced benefits like an increase in data quality, more comprehensive data analysis and an enhanced ownership of data at the lower levels in the field. It did require however a data model which was perceived as rigid to respond to dynamic needs and complex for data analysis in the field.

At the health facility level the actualization of functionalities was tightly coupled with an intuitive user experience that did not require training, and user motivation towards introducing the digital system in their day -to-day medical work was mainly based on the system providing such intuitive user experience and/or being explicitly useful for their daily responsibilities (Vila-Pozo et al., 2022). Under the operating conditions identified at the work-place in the humanitarian intervention, the way technology is perceived and more importantly, its perceived usefulness, influence how technology functionalities are realized. The analysis conducted in the study (Vila-Pozo et al., 2022) revealed that users consistently favoured technology solutions that were perceived as easy to use, requiring minimal training efforts and without any friction. However, this preference shifted when the technology demonstrated an immediate perception of usefulness, particularly through significant time savings in data management compared to traditional paper-based methods (Vila-Pozo et al., 2022).

Using Inductive Theorizing to Elaborate the Principles for Effective Use in the Context of Study

The main goal of the final step of this contextualization is to perform an analysis of the suitability of the theoretical model for effective use and its application in a humanitarian setting. To guide this analysis, I summarize the findings from the previous steps in three principles for effective use of information systems in humanitarian settings.

1. Humanitarian projects experience high turnover, and users work in hectic and autonomous environments. The completion of training tasks is challenging, and user adaptations may lead to system abandonment.
2. Humanitarian work develops in non-structured and fast-paced environments, an intuitive user experience is considered imperative to ensure use of the system.
3. Humanitarian workers have to juggle multiple priorities constantly, optimizing time and perceived usefulness are critical factors for effective system use within this context.

The theoretical model for effective use in a n emergency environment is depicted in Figure 4 below.

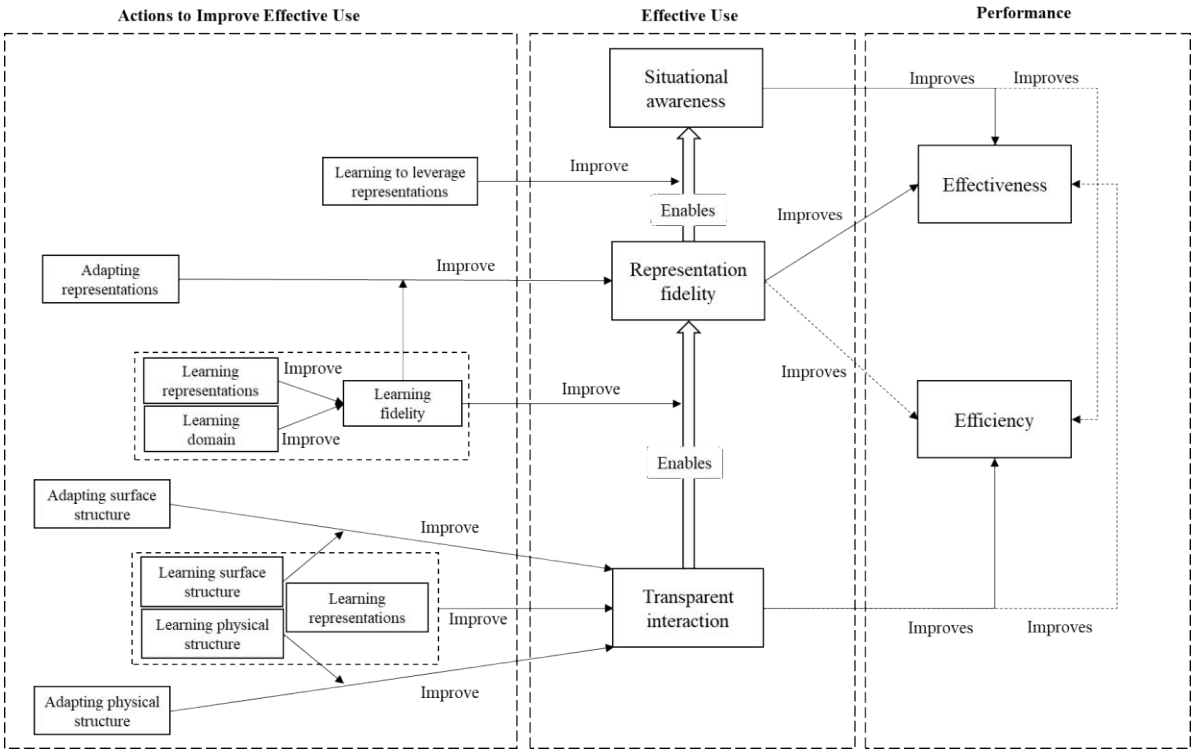


Figure 4. Representation of the TEU model (Burton-Jones & Grange, 2013) extended for emergency management (Bonaretti & Piccoli, 2018)

The image guides a discussion in which the fundamental principles of both the theory and the setting are assessed in combination. The analysis focuses on examining the dimensions for effective use, as defined in the original theory (Burton-Jones & Grange, 2013), which include transparent interaction, representational fidelity, and situational awareness (Bonaretti & Piccoli, 2018). The aim is to assess their suitability for the users and context of the H-HMIS, rather than customizing the dimensions specifically for this context, as suggested by the authors in their approach to contextualizing effective use (Burton-Jones & Volkoff, 2017) . The contextualisation approach taken in this study, hence, is based on the assessment of the nomological principles of the theory for effective use in contrast with the three principles for effective use in humanitarian settings derived from the previously presented findings.

The discussion will commence by understanding the relationship between the identified functionalities and the dimensions for effective use, followed by an assessment of the overall theoretical model in contrast with the three principles for effective use in humanitarian settings.

Effective Use through Effective Actualization

Situational awareness, as referenced in the literature (Endsley, 1995), is a precedent for informed action and is associated with understanding and comprehending the domain of work in context. This dimension plays a crucial role in enabling functionalities that involve decision-making at higher levels, including operationability, accountability, and contextuability in project coordination. Situational awareness is also relevant in tasks such as ward management and accessing patient data at the healthcare facilities. By enhancing the understanding of the situation, this dimension has the potential to enhance effectiveness and improve efficiency by helping guide the intervention providing information for informed decision-making, and minimizing the time required to rectify errors resulting from ill-informed decisions.

Representational fidelity refers to "the extent to which a user is obtaining representations from the system that faithfully reflect the domain being represented" (Burton-Jones & Grange, 2013). This dimension enhances effectiveness by facilitating users' comprehension of their domain and reducing uncertainty. And improves efficiency by reducing the time required to verify information that is mistrusted by the user. Representational fidelity enables situational awareness, thereby contributing to the functionalities associated with decision-making. In the context of this case study, it will also impact the functionality of collectability, as the limitations on available health related information in the system configuration can pose challenges to data collection when does not reflect properly the contextual needs. It will also influence the functionality of aggregability, as the subsequent

representations generated for data analysis are based on the periodically aggregated information.

Transparent interaction refers to "the extent to which a user is accessing the system's representations unimpeded by its surface and physical structures" (Burton-Jones & Grange, 2013). This dimension contributes to efficiency primarily by reducing the time users allocate to interacting with the system. Also, the system's capacity to provision useful representations can facilitate more seamless interactions, which, in turn, can enhance us er focus—a crucial factor for efficacy. This dimension is about system access. As such, it holds the initial position within the hierarchical construct of effective use, being placed at the point direct interaction occurs between the user and the technology. When analyzing technology use in the setting under study, the key conditions for the actualization of the functionalities required a seamless experience of system use and a direct perception of usefulness of the system as pivotal conditions for realizing the functionalities that support users' operational tasks.

The Model of Effective Use in the Humanitarian Environment

Actions to Improve Effective Use

The theoretical model suggests that users can do certain actions to improve how effectively they use the system. There are two types of actions: adaptation and learning. Learning actions are those taken to learn the system, the domain it represents or its representations. Adaptation actions are those made to improve the system representations or the access to them. According to the model (shown in Figure 4. Representation of the TEU model (Burton-Jones & Grange, 2013) extended for emergency management (Bonaretti & Piccoli, 2018)), these actions aim to improve the different dimensions of effective use, transparent interaction, representational fidelity, and situational awareness. They aim however more prominently at improving transparent interaction, which has a primary role on the dimension ladder (Burton-Jones & Grange, 2013).

When analysing the drivers for effective use, the first of the three principles for effective us in humanitarian contexts contrasts with the theory of effective use in the perception of the role of learning and adaptations. The theory of effective use positions learnings and adaptations as the mechanisms for improving the dimensions for effective use, while the first principle for effective use of information systems in humanitarian settings position these actions as a challenge and barrier to system action based on work environment and high turnover rates: "1. Humanitarian projects experience high turnover and users work in hectic and autonomous environments. The completion of training tasks is challenging, and user adaptations may lead to the abandonment of the system ".

This arguably constitutes a limitation in the applicability of EUT for humanitarian settings.

The Dimensions of Effective Use

Situational awareness is the dimensions which operates farther from the system -user interaction, as it is enacted through the processing of the information retrieved from the system to gain knowledge about a particular situation or aspect of a domain. In this sense, the principles for effective use in humanitarian settings do not question this construct, the opposite, they define the conditions the system should offer at the precedent dimensions to guarantee that users can attain situational awareness.

Representational fidelity. As stated in the third principle for effective use of information systems in humanitarian settings: "3. Humanitarian workers have to juggle multiple priorities constantly, optimizing time and perceived usefulness are critical factors for effective system use within this context". Meaning that users need to perceive that the system is useful. The representations provided play a key role in providing such experience as they need to inform about the situation the user is traying to understand and analyse is a simple way. Users need to read and process information in an agile and intuitive manner. If having good understanding of the representation requires training or adaptations, the system would be challenging the first principle, which state that "3. … the completion of training tasks is challenging, and user adaptations may lead to system abandonment and an intuitive user experience is considered imperative to ensure use of the system ".

Transparent interaction addresses system access as it is the "port of entry" for the interaction user-technology. The three principles for effective use pose challenges to this dimension, as it is a that level where the humanitarian setting imposes more restrictions. The more explicit mention to system action and user interaction is on the second principle which states that "2. …an intuitive user experience is considered imperative to ensure use of the system ". This couples with a need to minimize training, ideally system use would not require a cognitive process or explicit training because, according to principle "… the completion of training tasks is challenging" and would not require user adaptations as humanitarian workers operate in a very decentralized and autonomous way and "…user adaptations may lead to system abandonment". The last principle puts even more conditions to the user experience

as it adds a requirement of perceived usefulness. In an environment where "… workers have to juggle multiple priorities constantly", and the system must compete with other duties the perception of usefulness is a critical factor for effective system use.

Performance

The dimensions of effective use are necessary for improving performance through both effectiveness, which refers to "the extent to which a user has attained the goals of the task for which the system was used ", and efficiency, which relates to "the extent of goal attainment for a given input" (Burton-Jones & Grange, 2013). Performance is addressed from all the three principles for effective use in a humanitarian context but is the key focus of the third one stating that "3. Humanitarian workers need to constantly juggle priorities, optimizing time (efficiency) and ensuring perceived usefulness (efficacy) are critical factors for effective system use within this context."

In summary, the conceptualization of the dimensions for effective use as a hierarchical mechanism towards improving performance, aligns well with the contextual analysis of the H -HMIS and the individual functionalities identified in this contextualization. The analysis highlights, however, two main aspects where the theoretical model and principles might arguably not be directly applicable. Firstly, relying on user's actions, such as training and adaptation, to improve effective use is arguably not appropriate for the humanitarian context, and secondly, transparent interaction holds the majority of the challenges imposed by the context while being a determinant in system use and adoption of the H-HMIS under the conditions of the humanitarian environment.

These results raise an evident concern, humanitarian organisations need to improve transparent interaction while reducing the need for training and drastic adaptations. This builds the foundation to answer the research question which is addressed and discussed in the next section.

DISCUSSION

The contextualization of effective use in the case of a H-HMIS in a field intervention has provided the basis to answer the research question posed in this research *What are the challenges for attaining effective use of a humanitarian health management information system?*

In conclusion, the analysis and findings can be presented around two main aspects of effective use:

i) Transparent interaction is an essential dimension for effective use in humanitarian settings, being a major determinant in system use and adoption.
ii) Relying on user's actions (such as training and adaptation) to improve effective use through its dimensions, is arguably not appropriate for the humanitarian context.

These results position the challenges for attaining effective use of a humanitarian health management information system at the level of system access, with special emphasis in user experience and training actions as the ma in challenges for effective use is such work conditions. The role of transparent interaction is fundamental to address these concerns, hence this discussion will focus on understanding the construct and its application in the setting of study.

The Role of Transparent Interaction in Humanitarian Settings

The user experience of an H-HMIS needs to be extremely intuitive and provide a feeling of usefulness. In addition, due to contextual conditions, to complement an adequate level of transparent interaction, organizations need to minimize the need of training and adaptations. The main challenge to effective use is providing a maximized transparent interaction. This strong dependency leads to the refinement of the original definition in order to develop one that accounts for the particularities of the humanitarian setting.

Transparent is the adjective that describes the qualities of the user system interactio n. To refine the concept of transparent interaction, I will start by focusing on its etymological meaning, *Transparent*[2] means "*Having the property of transmitting light, so as to render bodies lying beyond completely visible; that can be seen through; diaphanous*". In figurative terms "*Easily seen through, recognized, understood, or detected; manifest, evident, obvious, clear*".

When applied to information system, this can be interpreted as the ability of the system to let the usefulness (as the light does in surfaces) pass through its structures. When this happens, the potential actions offered to the user

[2] Oxford English Dictionary

are perceivable in an intuitive manner. Users would see that they can perform certain tasks, for example, registering a patient, by looking at the user interface and interpreting its labels and menus. And that action will be important for them and contribute to their day-to-day tasks to a level that dedicating time to it compensates the required dedication of time and effort. This understanding of the interaction aligns the original definition of affordances as "the affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill" (Gibson, 1979). The conceptualization of affordances assumes no need of an explicit cognitive process for the animal to actualize the affordance, affordances are perceived. While the concept of affordances remains at the theoretical underpinning of th is research, the interaction of the user and the technology has been referred to as functionalities in this work. Hence, in the context of this article, this would be described as the intuitive perception and use of a functionality of the technology.

Winograd and Flores (Winograd & Flores, 1986) described readiness-to-hand as a very similar experience. They introduced it when describing transparent interaction in the domain of user experience and inspired the authors of the theory of effective use to propose the term *transparent interaction* as the first dimension (Burton-Jones & Grange, 2013). Readiness-to-hand emphasizes that users know the ultimate goal they want to achieve, such as driving a car, rather than the specific details of how to accomplish it, like determining the precise angle to turn the steering wheel. It highlights the user's perception of the possible actions involved (Winograd & Flores, 1986). This user experience relies on providing "the right coupling between the user and the action s/he wants to perform in the relevant domain" (Winograd & Flores, 1986). The domain refers to the reality in which the user develops its professional activities, and the key mechanism to coupling the system and the domain relies on the system´s ontology (Winograd & Flores, 1986). This emphasizes the importance of designing a system that uses a langua ge which aligns with the environment where it is used. This is particularly relevant for humanitarian interventions which develop a specific vocabulary sometimes based on operational terms or specific acronyms.

The original term "Transparent Interaction" is described in the EUT model as: the extent to which a user is accessing the system's representations not impeded by its surface and physical structures (Burton-Jones & Grange, 2013). Based on this analysis, this paper proposes to expand the definition of transparent interaction in two aspects: the system structures influencing transparent interaction and the understanding of what does it entail to impede system access.

In relation to the system structures, the definition of transparent interaction explicitly mentions the access through the surface and physical structures, not including the deep structure. The analysis of the operating conditions, however, revealed the deep structure as a fundamental influence in the final user experience. In some systems the deep structure defines the ontology of the system which shapes the surface structure. The available data points or variables shaping the analysis are defined at the deep structure and can impose limitations which come from the inner data model which and cannot be overcome by a user interface. The adequacy of the system definition with the domain of work is of utmost importance and an inadequate configuration can lead to low understanding of the system and a weakened sense of usefulness. The importance of the deep structure in creating the basis for an appropriate domain for the user is key for providing an intuitive and "transparent" interaction. The first modification proposed to the definition of Transparent interaction is to make it inclusive to the three levels of the information system structures.

In relation to system access, the proposed definition mentions "impeded access" but is not explicit about how access can be impeded, or what constitutes unimpeded access. Findings from previous research showed that functionalities offering natural interaction with the system and perceived intuitively were more likely to be actualized over other ones which required learning or bigger efforts (Vila-Pozo et al., 2022; Vila-Pozo & Sahay, 2019). If we combine that user experience with the fact that learning actions are not anadequate driver for effective use, then the need of learning itself becomes an impediment to both the intuitiveness and the perception of usefulness for specific goals.

Both, Gibson in its original definition of affordances, and Winograd and Flores in their understanding of readiness-to-use, point to the same type of friction-free experience required for actualizing functionalities in humanitarian settings. And bringing back the etymological meaning of transparency as the transparent material "allowing light to pass through so that objects behind can be distinctly seen", this work proposes that transparent interaction is attained when the structures provide a n extremely natural level of access in which the user can "see through" the system structures and perceive the required actions to attain a goal and its potential usefulness.

With these two considerations in mind, the proposed definition for transparent interaction is:

> *Transparent interaction*: the extent to which a user is accessing the system's representations in a way that its potential use is perceived through a surface, physical and deep structures.

This definition builds upon three qualitative field studies which scrutinized the organizational environment, the

functionalities, and their process of actualization of the H-HMIS in a humanitarian setting.

The definition and analysis of the concept of transparent interaction encapsulates two key aspects of the dimension: *(i)* the level of training required to use the system and *(ii)* the alignment between the system's ontology and the domain in which the user operates. Both aspects could be relevant in building measurement instruments for evaluating transparent interaction.

CONTRIBUTIONS AND FUTURE WORK

The contextualization of the theory of effective use in the humanitarian setting is accomplished through three means. Firstly, it involves studying the micro network of functionalities enabling effective use in the field settings. Secondly, it entails studying the micro-level process of actualization when situated in context. Lastly, it concludes by revisiting the dimension of transparent interaction and proposing two specific aspects for improving its measurement.

From a theoretical perspective, the analytical work contributes the theory of effective use by contextualizing and testing the theoretical model in the novel domain of humanitarian interventions, leading to the development of an empirically based definition of transparent interaction. Findings of this research identify learning and adaptations as a barrier for system use, and the development of the new definition for transparent interaction points out that coupling the system with the domain through an adequate and specific ontology is an integral mechanism for enhancingthe transparency of surface, physical, and deep structures to the user and reducing the need of training. Practically, this work aims at providing practitioners and researcher new paths for studying and strengthening health information systems in humanitarian field settings. The proposed definition can be instrumental for helping practitioners assess the level of transparent interaction of their systems identified as they key challenge for attaining effective use.

ACKNOWLEDGMENTS

I would like to explicitly express my recognition and gratitude to the organization, MSF Spain. Headquarters office, field mission and field project staff who made this research possible. I am deeply grateful and appreciate the efforts made and the time dedicated to contributing to this project amidst their already hectic schedules.

REFERENCES

ALNAP. (2015). *The State of The Humanitarian System*. A. Study.

ALNAP. (2022). *The State of The Humanitarian System*. A. Study.

Bonaretti, D., & Piccoli, G. (2018). Effective Use of Information Systems in Emergency Management: A Representation Theory Perspective. International Conference on Information Systems (ICIS 2018), San Francisco, CA, USA.

Burton-Jones, A., & Grange, C. (2013). From Use to Effective Use: A Representation Theory Perspective. *Information Systems Research*, *24*(3), 632-658. https://doi.org/10.1287/isre.1120.0444

Burton-Jones, A., Recker, J., Indulska, M., Green, P., & Weber, R. (2017). Assessing Representation Theory with a Framework for Pursuing Success and Failure. *MIS Quarterly*, *41*(4), 1307-1334. https://doi.org/10.25300/MISQ/2017/41.4.13

Burton-Jones, A., & Volkoff, O. (2017). How Can We Develop Contextualized Theories of Effective Use? A Demonstration in the Context of Community-Care Electronic Health Records. *Information Systems Research*, *28*(3), 468-489. https://doi.org/10.1287/isre.2017.0702

Checchi, F., Warsame, A., Treacy-Wong, V., Polonsky, J., Ommeren, M., & Prudhon, C. (2017). Public health information in crisis-affected populations: A review of methods and their use for advocacy and action. *The Lancet*, *390*(10109), 2297-2313. https://doi.org/10.1016/S0140-6736(17)30702-X

Eden, R., Akhlaghpour, S., Staib, A., & Sullivan, C. B.-J., Andrew. (2018). Unpacking the Complexity of Consistency: Insights from a Grounded Theory Study of the Effective Use of Electronic Medical Records. Hawaii International Conference on System Sciences, Hawaii.

Eden, R., Burton-Jones, A., & Donovan, R. (2019). Testing the Links from Fit to Effective Use to Impact: A Digital Hospital Case. International Conference on Information Systems (ICIS 2019), Munich, Germany.

Eden, R., Fielt, E., & Murphy, G. (2020). Advancing the Theory of Effective Use through Operationalization. European Conference on Information Systems (ECIS 2020), Virtual.

Endsley, M. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32-64. https://doi.org/10.1518/001872095779049543

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin.

Hunt, M., Pringle, J., Christen, M., Eckenwiler, L., Schwartz, L., & Davé, A. (2016). Ethics of emergent information and communication technology applications in humanitarian medical assistance. *International Health*, *8*(4), 239-245.

Haake, P. S., Silvia, Mueller, B., Lauterbach, J., & Maedche, A. (2018). Toward an Operationalization of Effective Use. European Conference on Information Systems (ECIS 2016), Portsmouth, United Kingdom.

Jarrett, P., Fozdar, Y., Abdelmagid, N., & Checchi, F. (2021). Healthcare governance during humanitarian responses: a survey of current practice among international humanitarian actors. *Conflict and Health*, *15*(1). https://doi.org/10.1186/s13031-021-00355-8

Magnuson, J. A., & Dixon, B. (2020). *Public Health Informatics and Information Systems*. Spinger Nature. https://doi.org/10.1007/978-3-030-41215-9

Mesmar, S., Talhouk, R., Akik, C., Olivier, P., Elhajj, I. H., Elbassuoni, S., Armoush, S., Kalot, J., Balaam, M., Germani, A., & Ghattas, H. (2016). The impact of digital technology on health of populations affected by humanitarian crises: Recent innovations and current gaps. *Journal of Public Health Policy*, *37*(Suppl 2), 167-200. https://doi.org/10.1057/s41271-016-0040-1

MSF Spain. (2021). *Informe de Operaciones*.

OCHA, U. (2022). *Global Humanitarian Overview 2023* (9210024133).

Strong, D. M., Volkoff, O., Johnson, S. A., Pelletier, L. R., tulu, B., Bar-On, I., Trudel, J., & Garber, L. (2014). A Theory of Organization-EHR Affordance Actualization. *Journal of the Association for Information Systems*, *15*(2), 53-85.

Trieu, V.-H. (2013, 01/01). Extending the theory of effective use: The impact of enterprise architecture maturity stages on the effective use of business intelligence systems. International Conference on Information Systems (ICIS 2013), Milan, Italy.

Vila-Pozo, M. M., Martinez, R., & Sahay, S. (2022). Competing Affordance Potencies Shaping Data Collection: Case of Humanitarian Organizations. International Conference on Information Systems (ICIS 2022), Copenhagen, Denmark.

Vila-Pozo, M. M., & Sahay, S. (2018). Humanitarian Health Information Systems: Different challenges and responses. European Conference on Information Systems (ECIS 2018), Portsmouth, England.

Vila-Pozo, M. M., & Sahay, S. (2019). Institutional shaping of affordances: Implications on information use in global humanitarian organizations. International Conference on Social Implications of Computers in Developing Countries (IFIP 9.4), Dar es Salaam, Tanzania.

Wand, Y., & Weber, R. (1995). On the deep structure of information systems. *Information Systems Journal*, *5*(3), 203-223.

Winograd, T., & Flores, F. (1986). *Understanding Computers and Cognition: A New Foundation for Design*. Intellect Books.

Avoiding "metal astray": experience evaluation of virtual reality training for melting furnace operators.

Jo Dugstad Wake*
NORCE
jowa@norceresearch.no

Frode Guribye
University of Bergen
frode.guribye@uib.no

ABSTRACT

The work of melting furnace operators is potentially hazardous and serious accidents do occur. Training on the monitoring and safe operation of furnaces is key to the safety of the operators and their surroundings. Virtual Reality (VR) environments can simulate safety critical work environments and provide realistic, safe and engaging ways to train on situations otherwise unavailable. To address the potential of using VR in safety training, the VR environment called VR Safe was developed. This paper presents a study of the user experience of training safety critical procedures in VR. VR Safe simulates malfunction in one of the most important systems regarding safe operation of furnaces: furnace cooling system monitoring and how to respond to cooling system failure. The evaluation included professional furnace operators, and was focused on usability of the system, user engagement and the training value as it relates to the operators' work.

Keywords

Safety critical training, virtual reality, usability, melting furnace operation.

INTRODUCTION, BACKGROUND AND RELATED WORK

Training is a common strategy to reduce the risk of accidents in hazardous work environments and is centered around increasing the skills of humans to identify, analyse and assess the risks involved in a particular environment or situation (Scorgie et al., 2024). Traditionally, training is facilitated by making use of different types of learning aids such as manuals, lectures, videos and drills (Feng et al., 2018). Managing high-risk environments requires a tight coupling between understanding unwanted situations as they unfold, being able to manually carry out the necessary operations, whilst simultaneously experiencing stress. Ultimately it is about developing competences in both knowing and doing. Virtual Reality (VR) environments is an emerging and promising platform for training in contexts such as these as it offers an engaging mode of interaction by immersing the user in a representation of a real environment that offers rich interaction modes (Grassini & Laumann, 2020), feedback in the form of observing the result of actions, the development of hands-on skills and learning-by-doing, and opportunities for repetition (Deb et al., 2017). Particularly relevant to training for high-risk situations, VR can also provide first-person access to environments that are potentially dangerous, unethical, costly, temporally unavailable or unpractical to access in the real world (Freina & Ott, 2015).

VR systems are being used in education and training for their potential for creating immersive experiences, in which the users can experience a sense of presence (Jensen & Konradsen, 2018). Examples of VR-training in safety critical environments are available for firefighting (Wijkmark & Heldal, 2020; Wijkmark et al., 2021); first responders (Binsch et al., 2021; Mossel et al., 2021) and several other areas. VR in safety critical training has also been studied in industrial work settings, for example working from heights (Cyma-Wejchenig et al., 2020; Loreto et al., 2018), working in confined spaces (Lu et al., 2020), overhead load movement and operation of cranes (Zhang et al., 2023) and electrical/high-voltage safety training (Ayala García et al., 2016; Wang & Messner, 2020). While VR environments for various purposes connected to metallurgical production and melting furnace operation exist, for example visualising and simulating processes (e.g. (Freitag & Urness, 2002; Fu et al., 2009; Ratts et al., 2000)),

*corresponding author

little work is done on VR-based safety critical training in this domain. Li and colleagues (Li et al., 2019) describe a system developed for safety training in metallurgical production, but their approach is more oriented towards technical systems description, and not on the experience of the operator.

This paper presents a user experience evaluation of a VR environment designed for safety critical training for operators of the hazardous work environment of melting furnaces in metallurgical production. The case illustrates the potential of using VR-training for safety critical and hazardous environments. In the study we recruited professional metal furnace operators to participate in the VR-training and focus on how they experience the training. The empirical data consists of physiological sensor data, answers to standardized questionnaires on usability, user experience and VR experience, and in-depth interviews with the operators. The data is used to explore the perceived usability of the environment, the way in which it engages the operators, and what kind of training experience the environment facilitates.

The study provides a rich account of how VR-training can be used in safety critical industrial work settings with a focus on how VR environments can provide realistic training situations. Further, we discuss how VR-training experiences can be evaluated by including different kinds of empirical data to address different aspects of the VR experience including engagement, usability and user experience, and the training value.

TECHNOLOGY AND CONTEXT: MELTING FURNACE OPERATION AND TRAINING

Melting Furnace Operation and Training

The educational requirement for the staff who operate the melting furnaces is a certificate of apprenticeship as industry process operator or chemical process operator. There are several safety risks connected to operating the furnaces, and the most commonly referred to hazard is called "metal astray". "Metal astray" means metal in liquid form anywhere else than where it is designed to be handled. "Metal astray" covers both minor or major spilling from containers, but also ejections from the furnace as a result from explosions. To avoid metal astray, the furnaces need to operate at the correct temperature. To ensure correct temperature, the furnaces rely on a cooling system consisting of a set of pipes streaming pressurised water surrounding the furnace exterior. Cooling water leakage occurs infrequently and episodically, but causes a situation that can be very dangerous because of the high energy matters and volumes involved. These situations require intervention by the operators, and how to intervene depends on the nature of the leakage, and the understanding of which depends on the perception, experience and discretion of the operator. The water pressure in the cooling pipes is monitored in a control room. Falling pressure, readable on manometer displays, is an indicator of leakage in a pipe, which reduces cooling. The leakage is not visible to the naked eye because of the temperature of the water and visual hindrances caused by the furnace (fire, smoke), but a mild reduction of the furnace temperature can make the leakage visible. These situations occur very rarely, but when they do the operator is usually alone. The risks are both economical and physical. Shutting down the furnace if not necessary is very expensive, whilst not shutting down if required is potentially catastrophic.

The current practice for safety training at the plant involved in this study consists of lectures and bi-annual certification tests presented as imaginary scenarios with challenges responded to in writing. They also use a digital reference manual, where all the relevant procedures are described. Additionally there is on-the-job training and knowledge exchange between novice and experienced operators. The opportunities for practical and physical experimentation are limited, due to the costs of taking the furnace out of production. Needless to say, there are also limits on the experimentation with the furnace because of the dangers involved with mixing high temperatures (ca 1600 degrees Celsius) and large quantities of melted metal, electricity and pressurised water. With VR it is possible to overcome some of these limitations. With VR, the physical environment can be realistically recreated, including the visual, three-dimensional appearance of the plant, the sound, and how the environment is manipulated in a dynamic way. VR can also provide opportunities to train on safety procedures without stopping the furnace.

VR Safe

VR Safe is a prototype of a training environment for melting furnace operators, currently containing one training scenario. It is developed by a consortium of VR software companies, in collaboration a Norwegian metallurgical production company. VR Safe allows operators to handle loss of cooling of a furnace. The cooling system consists of several water pipes, and currently VR Safe can simulate malfunction in one of those pipes. Additional common safety hazards for operating furnaces have been been placed in the environment, for example misplaced objects and hindrances that can cause safety issues. VR Safe is being created using the Unity engine to recreate the environment of an existing furnace, with adhering cooling system, building and control room. VR Safe is interacted with using a Head Mounted Display (HMD). Figure 1 illustrates clockwise from top left: a. The furnace, b. The control room, c. Break area and d. Water pipe panel with valves.



Figure 1. VR Safe environment

When opening VR Safe, the operator is placed in the melting plant control room. When the operator looks around, he/she can see the computer displays with readings of the pressure in the water pipes (b.), the inspection window to the furnace opening (a.). When turning around, the operators can see the rest of the control room with the kitchen area and electrical boxes (c.). There is a lever for reducing the energy in the furnace. The operator can also navigate to the valve panel area and close the correct valve (d.). Additionally, the operator can navigate to the furnace itself. The currently available scenarios are 1. falling pressure in one of the pipes, and 2. cluttered work spaces around the furnace area.

EVALUATION APPROACH AND METHODS

Evaluation Approach

This study examines the experience of interacting with VR Safe. Three aspects of the interaction is in focus, all considered central to the overall goal of facilitating critical safety training for furnace operators: 1. engagement, 2. usability and 3. training value. Each of the aspects were examined in its own way. The overall approach is illustrated in Figure 2

How users engage with training exercises can have an impact on the outcome (Xin, 2022). While different ways to measure and understand engagement have been suggested (Doherty & Doherty, 2019; Flobak, 2023), we focus on how we can use physiological data as an indicator of engagement. As Electrodermal activity (EDA) has been used as an indicator of mental effort (Agarwal et al., 2021; Pijeira-Díaz et al., 2018), we explore how such data, obtained with the biosensor Empatica E4, can be used to inform our study of the experience of VR-safety training for furnace operators. We assume a correlation between EDA values and an affective response - here understood as levels of

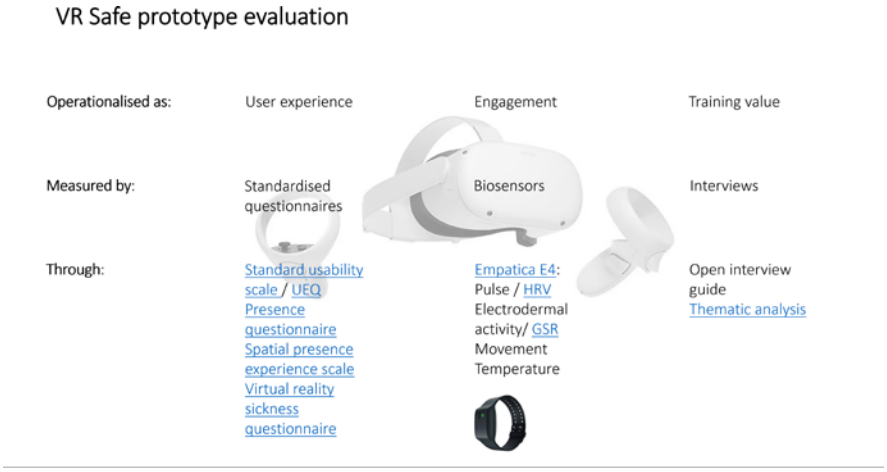


Figure 2. Overview of the evaluation

arousal, which again can be used as an indicator of the level of engagement with the environment and the training content.

The evaluation also focused on the user experience of the VR-training. In Human-Computer Interaction user experience is a term that can encompass both the usability of a system and the experience that the user has when using the system (Turner, 2023). We have approached the evaluation of the total VR-training experience as including both usability, user experience in general, but also including more specific elements that pertains to interaction in virtual environments.

The training value is seen in relation to the existing work and training practice for melting furnace operators. The training is studied in terms of how it relates to the operators established, professional understanding of work, with emphasis on critical safety procedures. These procedures consist of monitoring output from the furnace and the cooling system, and investigating and solving potential problems.

Methods

We collected physiological data for each participant to compare three different conditions: baseline, VR Safe use and VR gaming. Baseline was defined as taking part in a lecture for 15-20 minutes. Baseline should represent the "normal" training situation for the operators, but training takes place across a range of modalities as described above. We are interested in seeing how using VR Safe compares with taking part in lectures. Additionally, we included a VR game (Richies Plank Experience), to see how VR Safe compares with games that are designed for fun. Our expectation is that EDA will increase with VR Safe, as it involves the user in carrying out safety critical professional tasks, and expect it to increase even more when playing VR games. Physiological processes are individual, thus we are not interested in comparing persons, but rather how the different activities affect the physiological processes for an individual.

For this pilot study we decided to provide standardised questionnaires to expert/professional users. We chose to use questionnaires as we see it as an efficient way of getting user experience data with few participants accessible for a short time. The questionnaires are therefore analysed descriptively, and the results are seen as valuable when combined with the other data. We used two questionnaires for general usability assessment; System Usability Scale (SUS) (Brooke, 1996; Lewis, 2018), and User Experience Questionnaire (UEQ) (Schrepp et al., 2014). Two questionnaires were more specific to user experience with VR systems: Presence Questionnaire (PQ) (Witmer & Singer, 1998) and Spatial Presence Experience Scale (SPES) (Hartmann et al., 2015). Finally, two questionnaires for assessing any discomfort or unwanted bodily effects with use of VR Safe were used: Virtual Reality Sickness Questionnaire (VRSQ) (Kim et al., 2018) and Cyber Sickness Questionnaire – VR (CSQ-VR) (Kourtesis et al., 2023).

The part of the evaluation considering training value is about how the operators work in a wider context of colleagues, information systems, tasks, goals, tools and work culture. To access this information, an open qualitative interview guide was constructed. The themes included are how the operators understand learning goals for VR Safe,



Figure 3. Evaluation timeline

engagement, their perception of interacting with VR Safe, realism and fidelity, the relevance of the content and tasks, immersion, comfort and ergonomics in addition to how they would like to see VR Safe further developed.

The timeline for the evaluation activities is illustrated in Figure 3.

EVALUATION WITH MELTING FURNACE OPERATORS

The evaluation with melting furnace operators was carried out over two days at a smelting plant. The first day involved observation and getting an overview of the basics of melting furnace operation. The second day was used for evaluation with operators. Five melting furnace operators took part. Two worked at the furnace that VR Safe is a replica of, whilst three worked on a different furnace. Four were young, yet experienced operators (age < 30), and one was more senior (age > 50). One participant had used VR Safe previously. The evaluation was carried out in groups, following the structure 2+2+1. Two were first introduced to VR Safe through listening to a presentation of VR Safe for 15-20 minutes. This was used as baseline for the physiological data capture, with one of the participants. Then they tried out VR Safe using one HMD each for around 20 minutes, whilst the physiological response was recorded for the same participant. Afterwards, they played RPE, for up to five minutes. Finally, they completed the questionnaires and took part in an interview. We recorded the physiological response for three of the participants, as we only had one Empatica available.

Results

EDA: The physiological responses are presented in Table 1, for participant 1, 3 and 5, clockwise from top left. The EDA levels for VR Safe are higher than for baseline for the three participants. Although EDA levels are individual (Baker et al., 2010), the tendencies between activities are similar, and seem to support our expectation that EDA levels would increase with engaging with VR Safe compared to baseline. For the two participants who played RPE, they show a (relatively) flat line before peaking rapidly.

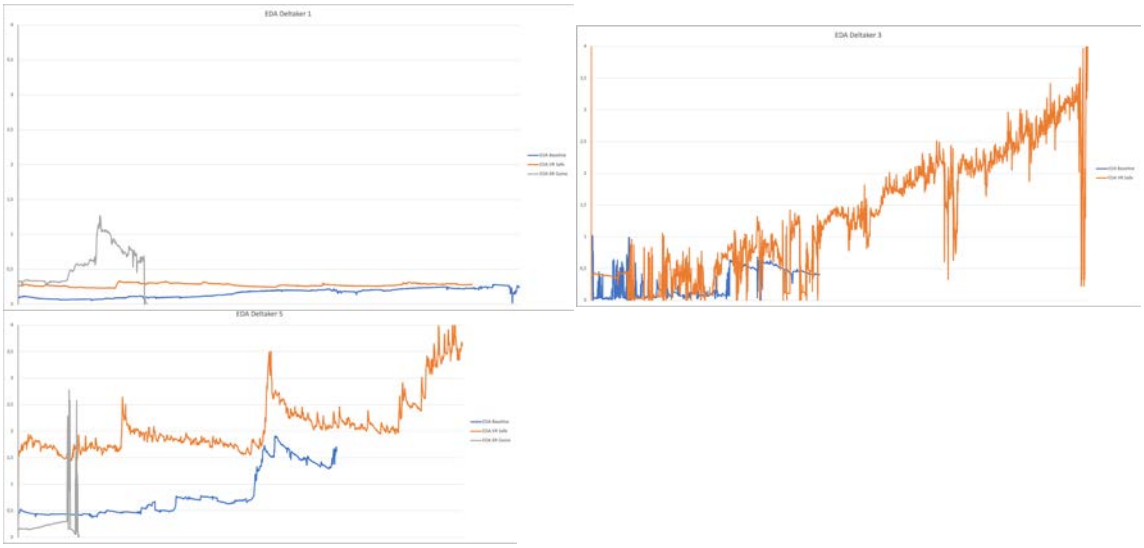


Table 1. EDA results, evaluation with operators

Participant 1 has marginally higher EDA levels with VR Safe compared to baseline, whilst both levels are low. RPE is higher, which increases trust in the data. The EDA levels for Participant 3 is higher than for baseline, and also increases noticeably throughout the activity. Participant 3 declined the RPE activity. The EDA levels for Participant 5 are higher with VR Safe than for baseline. The EDA levels for RPE seem to follow the same structure as for Participant 1. Overall, the EDA results can be an indication that the mental effort exerted with VR Safe is higher

than listening to a presentation and taking part in a discussion, and that training with VR Safe is a more active way of training.

Questionnaires: The results from the questionnaires have been visualised and analysed descriptively. In the visualisations, the positive/negative phrasing of questions have been rearranged to show the most positive responses to the right, to increase readability. For the two questionnaires measuring unwanted effects or discomfort following use of VR (the last two), the most positive responses are to the left.

The SUS questionnaire contains 10 statements about the perceived usability of a system. In the visualisation in Table 2, the SUS results appear mostly positive, with the most or second most affirmative responses being most frequent. The average SUS score is 77,5. For SUS scores, results over 70 are considered good. UEQ measures the qualities attractiveness, perspicuity, efficiency, dependability, stimulation and novelty. The major tendency in the replies are positive.

The results from the questionnaires that measure the VR experience are visualised in Table 3. The results are mainly positive, with the most or second most affirmative answer option being used for most of the items. The PQ contains how control and realism factors contribute to the experience of presence in a VR system (Witmer & Singer, 1998). The SPES questionnaire contains 20 items that measure the subjective feeling of being present in a virtual environment (Hartmann et al., 2015), and is based on the theory that to experience spatial presence, users first develop a mental model of the VR environment, and secondly accepts it as their egocentric viewpoint (ibid.). The users ability to visually create a mental model of the VR space affects the experience of spatial presence positively. This experience is relevant for the participants in our study, as the virtual environment is based on an environment that the participants know well.

The results from the two questionnaires measuring discomfort resulting from using VR are available in Table 4, where the lowest symptom reports are presented to the left in both tables. We consider these to indicate no serious discomfort from using VR Safe.

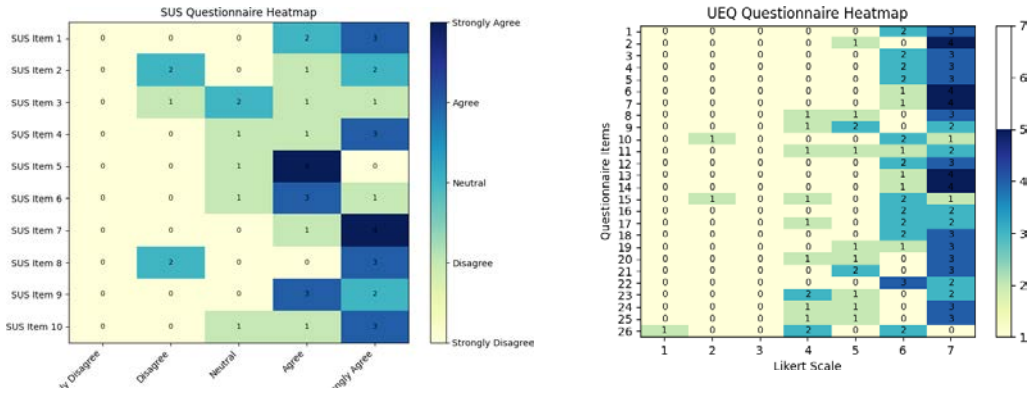


Table 2. Usability questionnaires

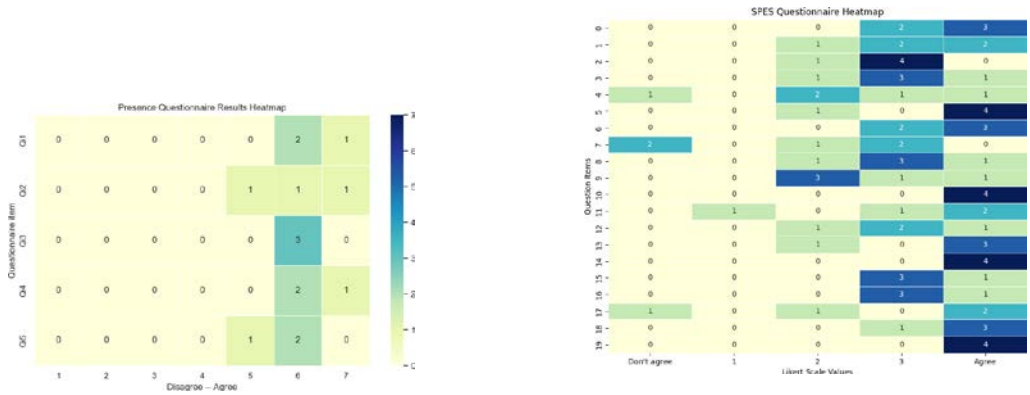


Table 3. VR experience questionnaires

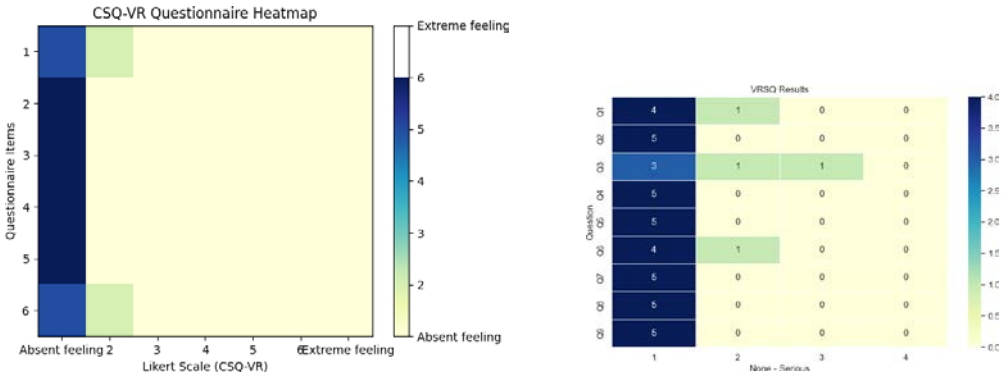


Table 4. Unwanted effects of VR use

Interviews: The interviews were recorded and transcribed in full before being analysed using Thematic analysis (Braun & Clarke, 2006). One theme is how the VR training is relevant to the existing training practices and needs at the smelting plant. This means in how it fits in with current practices, and whether and in which way it creates new training opportunities. The following two quotes illustrates how the operators considered the training relevant; it makes available situations that are unpractical, uneconomical or unsafe with their training practices today. There is also a need to train on practical intervention and solving problems in a manual way.

Participant 5: "I have trained a lot of new people (*operators*). And what is difficult with training new employees is, what do you do with a water leak? We can´t just go up and close a valve... "

Participant 5: "There is a lot of theory. But if you are unlucky and actually get a water leak, then you have to go up and close a valve. But here (*in VR Safe*) you actually get to do it. We are actually trying to do those things."

The operator is speaking about the notion of opening and closing furnace cooling valves on a furnace in production. The environment in which the operators work can be considered "low risk / high consequences". The low risk is contingent on following the designed procedures for handling high energy matter, and also being competent to perceive and intervene in randomly occurring unwanted situations. Tied to the theme of relevance is also managing the potential magnitude of the consequences, whilst experiencing stress. The following dialogue captures how they consider the consequences if things should go wrong:

Participant 1: "Yes, the potential that it can go wrong is very large, in case there should be an accident. I think that would be the very worst that could happen."
Participant 2: "Yes. That would be a crisis."
Participant 1: "Yes, you don´t want that to happen on your watch, to put it like that."

When considered as to how VR Safe is relevant to the safety training of the operators, the dialogue further highlights how being able to train repeatedly on detecting unfolding, unwanted situations and practical intervention becomes available in VR. VR also supports the formation of automated actions through repetition, as automated actions are much easier and probable to be undertaken when a person is experiencing a high state of stress.

Regarding the design of VR Safe, and whether it is developed in a way that the operators can recognise and find meaningful, the operator reported that they found it to be a useful way to carry out training. In the first quote, Participant 1 explains this by making a reference to the real world functioning of the furnace, and that VR Safe matches this. His conception of VR Safe is similar to that of the real life furnace. This is an indication of the importance of creating simulations that align well with real-world processes, when creating VR-based safety critical training environments. The operators who did not work at "Furnace 5" responded similarly to the operators that did, indicating that the formation of mental models corresponds to more than visual similarity of real and virtual environments, and also include dynamic relations in processes. This relates to the underlying theoretical foundations of spatial presence. Participant 3 highlights how VR-based training is more interesting than some other alternatives, such as the digital reference manual where every relevant procedure is described in detail in text.

Interviewer: "Was it an adequate way to learn to handle leakages in the cooling system?"
Participant 2: "Yes it is."
Participant 1: "It is sort of made the way that it actually works. It is quite easy to discover what you should do."

Participant 3: "Yes, it is, absolutely. I would rather use this than "Rosa". "Rosa" is where we keep all in-depth content. It is a system where we read and sign that we have understood (the procedures). It describes how to do everything in detail."

Regarding the visual representation of the furnace in VR Safe, it is considered to be adequate by the operators. One operator described the audiovisual appearance as meeting his expectations. "Participant 2: It appeared the way I imagined it would. The representation of objects..."

Participant 1: "The recognition effect. It is the way it looks."

Four of the participants had never used VR Safe previously. When asked about how it was to learn to interact with VR Safe, all participants reported that it was easy to use, although it took a little while to learn how to move for some (VR Safe uses the "teleport" mode for moving in the environment, where the user points the controller to a spot on the floor and click the joystick to move, or press the joystick forward to move in steps.). Once they had learned it, all participants found it easy to interact with VR Safe. One exception was operating the valves. The design for user input on objects is done by highlighting the objects when moving the hand/controller near the object. When highlighted, the valves can be opened/closed by making a twisting motion with the hand. This operation was experienced as difficult to perform by most operators. It is a central operation in connection with the safety procedures.

Participant 3: "(VR Safe) was easy to learn. That went smoothly. The only thing that did not work properly, was opening the valves."
Participant 4: "It worked okay to close them, but not opening them again."

DISCUSSION

In this study, we have engaged professional experts in testing a virtual environment for critical safety training. VR Safe simulates certain safety operations in the handling of a real furnace, and is, as such, a representation of key aspects of a real system for metallurgical production. To be useful for the operators in training, the virtual environment needs to simulate the real environment in a realistic way. To further understand the realism and perceived authenticity of the VR system, the study highlights how the fidelity of the representation affects the user experience. We describe these as **audiovisual, task** and **operational** fidelity. The audiovisual fidelity of representation concerns how the virtual environment matches the environment known to the operators. The task fidelity describes how the tasks in the virtual environment matches the tasks carried out by the operators while they work. The operational fidelity describes how the manual operations that the operators carry out in their work matches those in real life. In our study, the operators described the virtual environment as familiar to them in a visual sense, and that there was a recognition effect. The interviews also highlight that task representation is likely to affect their overall experience, including the audiovisual experience. The audiovisual representation in VR Safe contains a limited selection of the smelting plant, and even includes some minor mismatches with the real environment, such as it being possible to walk near the furnace (something that they would not to in the real environment, but instead use a forklift for protection). The fact that they understand the task (possible cooling malfunction), prompts them to carry out operations that they can understand based on their professional experience. The task fidelity, which speaks to the relevance of a task in the virtual environment, is likely to support the operators in forming a mental model of the virtual environment, and experience spatial presence. The interviews also revealed that there was an interactional problem with carrying out an operation important to the task, which was opening/closing the valves. This is a critical safety operation, needed for the safe operation of the furnace, and was not represented very well in VR Safe as it is a simple operation to complete on an actual furnace. Further refinement of VR Safe should look into different solutions for this operation. This identified shortcoming, points another kind of fidelity, namely haptic or tangible fidelity. To train on the physical and motor skills required for operating the system, can be hard to replicate in a virtual environment. Still, advances in haptic interaction in VR systems can possibly help to increase also the haptic fidelity and increase the realism of performing such tasks virtually.

The measurement of physiological responses of the participants revealed an increase in EDA levels whilst using VR Safe, indicating an increase in mental effort or activeness whilst using the VR environment. Activeness, and cognitive and emotional arousal has previously been found to be conducive to learning (see e.g. (Baker et al., 2010; Bjork et al., 2013; Pijeira-Díaz et al., 2018). Whilst the interview and questionnaire data are based on self reporting, the affective sensing represent more direct measurement of the process of engaging with VR Safe as it unfolds. Yet, we don't know exactly what in the VR experience that caused changes EDA levels. A further refinement of the method would be to connect physiological responses with what the participant is doing in VR at each moment, for example by streaming the view of the user. This approach could provide useful design information, for example by identifying particular tasks or other design elements that correlate with particular responses. The EDA data, and

particularly the VR Gaming data indicate that these kind of responses exist. Noori and colleagues (Noori et al., 2019), have demonstrated that continuous capture of physiological responses can be used to create VR systems that adapt to the users physiological responses for VR-based exposure therapy. In the context of safety critical training the responses could be used to increase or decrease difficulty or complexity of the task.

Considering VR Safe as a Proof-of-concept, the questionnaire data from expert participants seem to indicate that there are no major usability issues with VR Safe, and that the usability does not provide major hindrances to engaging in training based on VR.

Limitations: A major limitation in our study is that the data are based on a limited number of participants for the findings to be generalised to a larger population. Still we have recruited expert professional operators and there is a limited number of such professionals available. This limitation could be met by doing a less episodic study over time with several iterations of design and evaluation with professionals.

CONCLUSION AND NEXT STEPS

The paper has reported on a user experience study of professional operators engaged in VR-based safety critical training on handling malfunction of a melting furnace. A method for evaluating the experience focused on engagement, usability and training value was designed, and carried out in practice. We found the prototype VR Safe to be a method for safety critical training that has a promising potential as a way to overcome the limitations associated with other ways of training, such as the need for tying theoretical knowledge with operational practice and engaging the user.

Acknowledgements: Thanks to Gunnar Maltsev Hansen and Isabella Mikkelsen and the furnace operators at Elkem Bremanger for generously facilitating and participating in the evaluation. Also, thanks to Broomstick AS and Agder XR Productions for enabling the screenshots from VR Safe in Figure 1.

REFERENCES

Agarwal, A., Graft, J., Schroeder, N., & Romine, W. (2021). Sensor-Based Prediction of Mental Effort during Learning from Physiological Data: A Longitudinal Case Study. *Signals*, *2*(4), 886–901. https://doi.org/10.3390/signals2040051

Ayala García, A., Galván Bobadilla, I., & Arroyo Figueroa, G. (2016). Virtual reality training system for maintenance and operation of high-voltage overhead power lines. *Virtual Reality*, *20*, 27–40. https://doi.org/https://doi.org/10.1007/s10055-015-0280-6

Baker, S., D ́Mello, S., Rodrigo, M., & Graesser, A. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223–241. https://doi.org/https://doi.org/10.1016/j.ijhcs.2009.12.003

Binsch, O., Bottenheft, C., Landman, A., Roijendijk, L., & Vermetten, E. H. (2021). Testing the applicability of a virtual reality simulation platform for stress training of first responders. *Military Psychology*, *33*(3), 182–196. https://doi.org/10.1080/08995605.2021.1897494

Bjork, R., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*, *64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. *Usability evaluation in industry*, *189*(3), 189–194.

Cyma-Wejchenig, M., Tarnas, J., Marciniak, K., & Stemplewski, R. (2020). The influence of proprioceptive training with the use of virtual reality on postural stability of workers working at height. *Sensors (Switzerland)*, *20*(13), 1–17. https://doi.org/10.3390/s20133731

Deb, S., Carruth, D., Sween, R., Strawderman, L., & Garrison, T. (2017). Efficacy of virtual reality in pedestrian safety research. *Applied Ergonomics*, *65*, 449–460. https://doi.org/https://doi.org/10.1016/j.apergo.2017.03.007

Doherty, K., & Doherty, G. (2019). Engagement in HCI: Conception, Theory and Measurement. *ACM Computing Surveys*, *51*(5), 1–39. https://doi.org/https://doi.org/10.1145/3234149

Feng, Z., Gonzalez, V., Amor, R., Lovreglio, R., & Cbrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers in Education*, *127*, 252–266. https://doi.org/10.1016/j.compedu.2018.09.002

Flobak, E. (2023). *Meaningful Engagement with Digital Health Interventions : Experience, Participation and Design* [Doctoral dissertation, University of Bergen].

Freina, L., & Ott, M. (2015). A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives. *The international scientific conference elearning and software for education*, 10–1007.

Freitag, L., & Urness, T. (2002). Analyzing industrial furnace efficiency using comparative visualization in a virtual reality environment. *Journal of Manufacturing and Engineering*, *124*(2), 456–461.

Fu, D., Wu, B., Moreland, J., Chen, G., Ren, S., & Zhou, C. (2009). Virtual reality visualization of typical processes in blast furnace. *AISTech 2009-Proceedings of the Iron and Steel Technology Conference*, 307–314.

Grassini, S., & Laumann, K. (2020). Evaluating the use of virtual reality in work safety: A literature review. *30th European Safety and Reliability Conference, ESREL 2020 and 15th Probabilistic Safety Assessment and Management Conference, PSAM 2020*, 4964–4971.

Hartmann, T., Wirth, W., Schramm, H., Klimmt, C., Vorderer, P., Gysbers, A., Böcking, S., Ravaja, N., Laarni, J., Saari, T., & Gouveia, F. (2015). The spatial presence experience scale (SPES). *Journal of Media Psychology.*, *28*(1), 1–15. https://doi.org/https://doi.org/10.1027/1864-1105/a000137

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0

Kim, H., Park, J., Choi, Y., & Choe, M. (2018). Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics*, *69*, 66–73.

Kourtesis, P., Linnell, J., Amir, R., Argelaguet, F., & MacPherson, S. E. (2023). Cybersickness in Virtual Reality Questionnaire (CSQ-VR): A Validation and Comparison against SSQ and VRSQ. *Virtual Worlds*, *2*(1), 16–35. https://doi.org/10.3390/virtualworlds2010002

Lewis, J. (2018). The System Usability Scale: Past, Present, and Future. *International Journal of Human–Computer Interaction*, *34*(7), 577–590. https://doi.org/https://doi.org/10.1080/10447318.2018.1455307

Li, J., Sun, E., Li, Z., & Gao, D. (2019). Metallurgical Operation Safety Training System Based on 5G and VR. *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2204–2208.

Loreto, C. D., Chardonnet, J.-r., Ryard, J., Rousseau, A., Loreto, C. D., Chardonnet, J.-r., Ryard, J., Rousseau, A., & Virtual, W. A. (2018). WoaH : A Virtual Reality Work-at-Height Simulator. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces*, 281–288. https://doi.org/10.1109/VR.2018.8448292

Lu, S., Xu, W., Wang, F., Li, X., & Yang, J. (2020). Serious Game: Confined Space Rescue Based on Virtual Reality Technology. *2020 2nd International Conference on Video, Signal and Image Processing (VSIP '20)*, 66–73. https://doi.org/10.1145/3442705.3442716

Mossel, A., Schoenauer, C., Froeschl, M., Peer, A., Goellner, J., & Kaufmann, H. (2021). Immersive training of first responder squad leaders in untethered virtual reality. *Virtual Reality*, *25*(3), 745–759. https://doi.org/10.1007/s10055-020-00487-x

Noori, F. M., Kahlon, S., Lindner, P., Nordgreen, T., Tørresen, J., & Riegler, M. (2019). Heart Rate Prediction from Head Movement during Virtual Reality Treatment for Social Anxiety. *International Conference on Content-Based Multimedia Indexing (CBMI)*, 1–5. https://doi.org/10.1109/CBMI.2019.8877454.

Pijeira-Díaz, H. J., Drachsler, H., Kirschner, P. A., & Järvelä, S. (2018). Profiling sympathetic arousal in a physics course: How active are students? *Journal of Computer Assisted Learning*, *34*(4), 397–408. https://doi.org/10.1111/jcal.12271

Ratts, E., Murphey, Y., & Zhou, Y. (2000). Thermal modeling of controlled atmosphere brazing process using virtual reality technology. *Applied thermal engineering*, *20*(17), 1667–1678.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience: Third International Conference*, 383–392.

Scorgie, D., Feng, Z., Paes, D., Parisi, F., Yiu, T. W., & Lovreglio, R. (2024). Virtual reality for safety training: A systematic literature review and meta-analysis. *Safety Science*, *171*, e106372. https://doi.org/https://doi.org/10.1016/j.ssci.2023.106372

Turner, P. (2023). Creating a Good Experience. In *A psychology of user experience. human–computer interaction series.* (2nd ed., pp. 25–44). Springer Cham. https://doi.org/https://doi.org/10.1007/978-3-031-32454-3{_}2

Wang, X., & Messner, J. (2020). The Pedagogical Value of Virtual Reality Training for Electrical Workers on Energy Storage and Microgrid Systems. *Construction Research Congress 2020: Safety, Workforce, and Education*.

Wijkmark, C. H., & Heldal, I. (2020). Virtual and Live Simulation-Based Training for Incident Commanders. In A. L. Hughes, F. McNeill, & C. Zobel (Eds.), *Proceedings of the 17th iscram conference* (pp. 1154–1162).

Wijkmark, C. H., Heldal, I., & Log, M. M. M. (2021). Experiencing immersive VR simulation for firefighter skills training. In A. Adrot, R. Grace, K. Moore, & C. Zobel (Eds.), *Proceedings of the 18th iscram conference* (pp. 913–921).

Witmer, B., & Singer, M. (1998). Measuring Presence in Virtual Environments. *Presence*, *7*(3), 225–240.

Xin, Y. (2022). Influence of Learning Engagement on Learning Effect under a Virtual Reality (VR) Environment. *International Journal of Emerging Technologies in Learning (Online)*, *17*(5), 226–237. https://doi.org/10.3991/ijet.v17i05.29451

Zhang, Z., Wong, M., & Pan, W. (2023). Virtual reality enhanced multi-role collaboration in crane-lift training for modular construction. *Automation in Construction*, *150*, e104848. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104848

First Evaluation of the Usability and Usefulness of the PROACTIVE App

Laura Petersen
UIC
petersen@uic.org

Natasha McCrone
Rinisoft
natasha@rinicom.com

Grigore Havârneanu
UIC
havarneanu@uic.org

Garik Markarian
Rinisoft
garik@rinicom.com

ABSTRACT

One way to reduce the communication challenges associated with Chemical, Biological, Radiological, Nuclear and explosive (CBRNe) incidents management could be a public facing disaster smartphone app, of which none are currently available. The EU funded project PROACTIVE used an iterative, co-creation methodology to develop such an app. This work-in-progress paper examines the preliminary results of a questionnaire given to observers who were tasked with using the beta app during a first field training exercise. The results show that the app was found to be usable and useful. Based on the results of the current evaluation exercise, the app will be updated, taking into account the end-user inputs. Following which, the app will undergo testing in two additional field training exercises, in an iterative manner, before reaching the final, end-of-project version.

Keywords

CBRNe, Universal Design, disaster apps, co-creation, user interface.

INTRODUCTION

Chemical, Biological, Radiological, Nuclear and explosive (CBRNe) incidents present unique challenges for disaster management and crisis communication. For example, first responders may need to wear Personal Protective Equipment (PPE), which includes face masks, when interacting with victims and victims will most likely undergo unique protocols such as decontamination. Furthermore, while all disasters can lead to anxiety, seeing practitioners arrive in PPE on the scene while waiting to begin decontamination has been found to induce fear (Carbon et al., 2021).

These challenges are only compounded by the fact that public awareness of CBRNe incidents is low, which means that the public at large is likely to depend even more than in other types of disasters on the communication and technology employed by practitioners (Hall et al., 2020). One way to overcome or at least reduce these communication challenges as well as the public's anxiety about such events could be a dedicated CBRNe disaster app. Disaster apps are a widely recognized, useful tool for disaster management, contributing to the reduction of anxiety (Bossu et al., 2018), yet few available apps address the CBRNe topic (Petersen et al., 2021).

That said, it is important to keep in mind that the public is not a homogenous mass and a switch must be made from interacting with the public on a 'one-size-fits-all' basis towards more nuanced planning and preparing to deal with the range of individuals that together constitute a diverse 'public', e.g., disabled persons, older persons (Petersen et al., 2019). Unfortunately, rarely are digital tools for emergency management, including public facing disaster apps, designed with accessibility in mind (Gjøsæter et al., 2021).

Accessibility is a key component of social sustainability and inclusiveness, contributing to usability and durability – as stated in the European Accessibility Act (European Commission, 2019). It may be argued that accessibility

should have a key place in disaster apps, as when a disaster occurs, anxiety is likely to be high and therefore comprehension levels reduced. This is an effect known as situational disability , as indeed circumstances surrounding a disaster may lead to temporary impairments (Gjøsæter et al., 2019). One way to help make disaster apps accessible is to include universal design principles from the beginning of ICT (Information and Communication Technology) tool development (Radianti et al., 2017). Universal design is defined as the design of products and environments to be usable to the greatest extent possible by people of all ages and abilities. It is based on the following seven principles as elaborated by Connell et al. (1997): equitable use; flexibility in use; simple and intuitive use; perceptible information; tolerance for error; low physical effort; and size and space for approach and use.

The EU H2020 project PROACTIVE (PReparedness against CBRNE threats through cOmmon Approaches between security pra CTItioners and the VulnerablE civil society) presents a novel solution in the form of a mobile app to i) provide a public-facing CBRNe disaster app while ii) minimising the accessibility gap in such kinds of apps.

THE PROACTIVE MOBILE APP

The PROACTIVE mobile app was developed following the above-mentioned principles of universal design, based on input from the project's desk research phase, the experience of the app developers and a co-creation methodology, whereby both CBRNe practitioners and the civil society beta-tested a web-version of the app, as described in Petersen et al. (2021, 2022, 2023).

The app aims to improve the efficiency of communication between Law Enforcement Agencies (LEAs) and citizens around CBRNe incident preparedness and response by allowing LEAs to inform the public about ongoing incidents, to allow the public who have registered as app users to report to LEAs, and by providing CBRNe preparedness information materials to anyone who has the app.

Therefore, the app administrator is intended to be the LEA who would be responsible for CBRNe incident response and the main user of the app is intended to be the public at large. The app administrator/LEA would send the in-app notifications, create the incidents viewable on the app, receive the reported incidents from the registered public users and choose which information materials to include in the app repository.

The beta version of the mobile app (Figure 1) has the following four main features:

(i) In-app notifications: allow the administrator to push information about an ongoing incident. They are accessed by clicking on the "news" button at the bottom of the screen (Figure 1B).

(ii) Incident List: this section of the app shows the geographical location of all the incidents added to the app by the app administrator. The app administrator may include the following information about an incident: time, date, location (via GPS or manually inserted), classification (emergency, non-emergency), status (ongoing, awaiting information, resolved), headline, description and choose whether or not to send out an in-app notification.

The incident list is accessed by clicking on the "map" button at the button of the screen and a given incident can be seen in Figure 1D. Incidents are presented on a scrollable map and, below, a short description can be seen. More information can be found for each incident by clicking on either the dot on the map or the incident on the list.

(iii) Report an Incident: allows registered public users to provide information (e.g., time, location (via GPS or manually inserted), photos, description) about a suspected CBRNe incident. This feature is intended to aid CBRNe practitioners in evidence and information gathering (Gavel et al. 2021). It is accessed by clicking on the red button seen on the home screen (Figure 1A), and the screen to actually report the incident can be seen in Figure 1 C.

(iv) CBRNe Information: is a repository of public pre-incident information materials which are downloadable as PDFs. It can be accessed by clicking on the blue button on the home screen (Figure 1 A) and looks like Figure1E.

METHOD

This paper examines the preliminary results of a questionnaire on the usability and usefulness of the beta version of the PROACTIVE mobile app which was first tested during a field training exercise by observers.

Observers

The ISO 22398:2013 Societal security — Guidelines for exercises (ISO, 2013) defines an observer as an "exercise

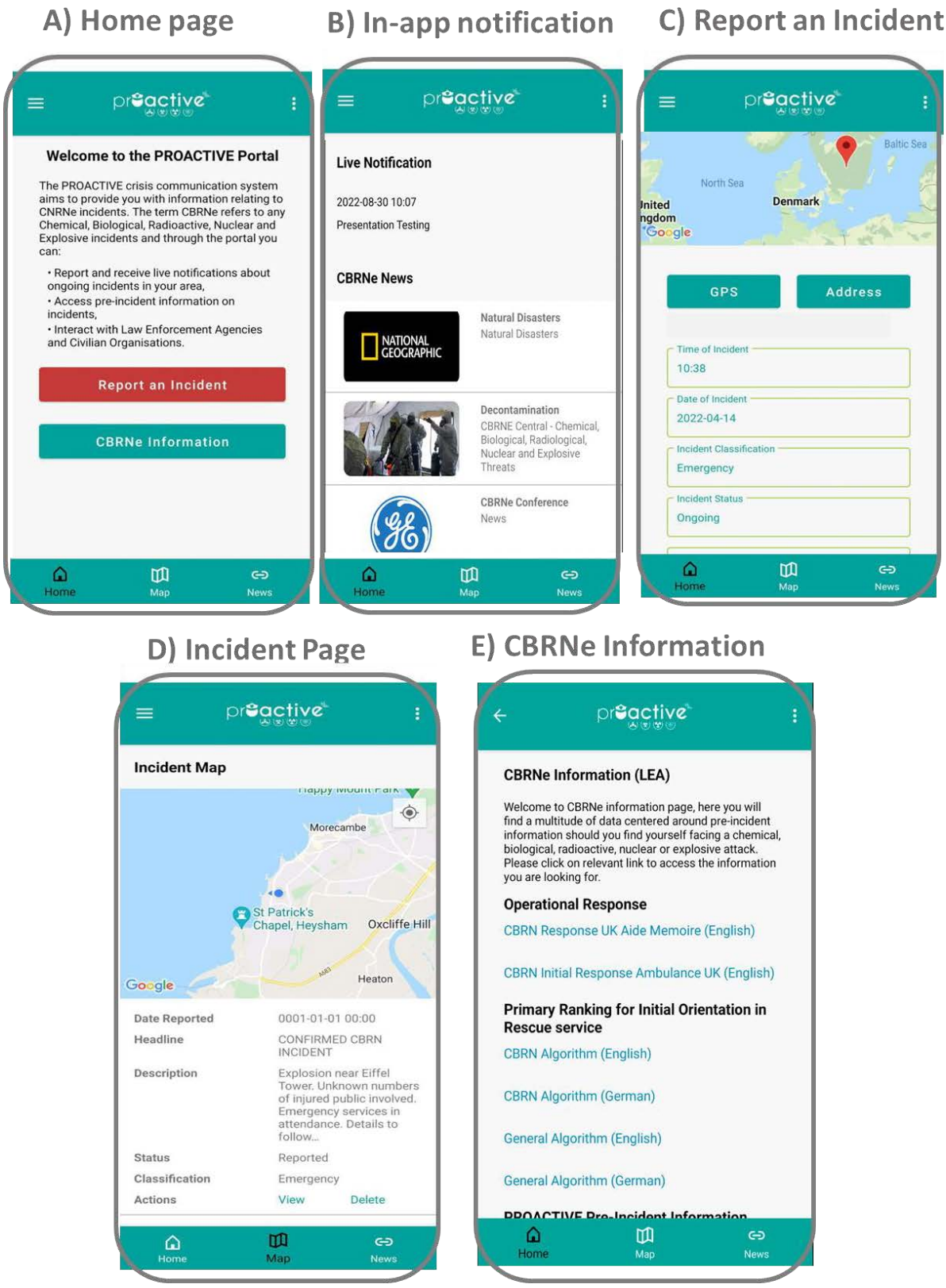


Figure 1. Graphic user interface (GUI) overview of the PROACTIVE app

participant who witnesses the exercise while remaining separate from exercise activities." Traditionally, the people invited to act as observers of a given field training exercise are security practitioners (e.g., firefighters, law enforcement, civil protection) with good knowledge of expected response protocols and standard operating procedures. A unique aspect of the PROACTIVE project was to also invite representatives of civil society organisations to attend the exercise as observers, to better understand the needs and expectations of the public during such incidents. As such, observers for the PROACTIVE field training exercise were dra wn from the two advisory boards of the PROACTIVE project: the Practitioner Stakeholder Advisory Board, made up of CBRNe practitioners, and the Civil Society Advisory Board, made up of civil society organisations.

During the field exercise one of the roles assigned to observers was to role play a witness to the incident who would use the PROACTIVE app as a registered, public user to a) look for information about the ongoing incident and b) use the 'report an incident' feature. In total, 33 observers were present for the field training exercise, of which 19 filled-in the questionnaire.

Field Training Exercise Scenario

The field training exercise took place on 7 May 2022 in Dortmund, Germany and was hosted by the Fire Department in Dortmund (FDDO) at their Civil Training Centre. The exercise was based on a scenario of a chemical substance from a freight train being released near a train station, leading to a specialised operational response. The chemical leak was simulated using disco fog in order to increase realism. The exercise was run in German and focused on the decontamination phase of emergency response. FDDO involved 130 firefighters who trained their decontamination skills. The role of victim was played by members of the local community, as opposed to the usual practice of off-duty first responders. The volunteers underwent disrobing, showering and rerobing.

During the field training exercise, seven pre-defined in-app notifications were sent out by the PROACTIVE app administrator, simulating notifications that would be sent by a LEA in the case of a real incident (Table 1). Each notification was sent in both English and German. Furthermore, the app was translated into German and available in both languages during the exercise.

Time	Notification
Exercise start	It is confirmed that an incident occurred around 07.00 this morning on the 7 th May at the local railway station. Reports have been received of smoke near a train. Please avoid the area as much as possible. Further information is being collected, please check this App for further updates.
+45 minutes	Confirmation received of a chemical substance leak requiring full decontamination. All people involved in the incident are being asked to remain calm, stay in the designated area indicated by the Fire Brigade and wait for instructions.
+1 hour	People requiring First Aid are asked to make the authorities on site aware of their needs
+1 hour 15 minutes	First responders have arrived on scene. A decontamination procedure is underway. Please keep calm and follow the instructions provided by the Fire Brigade on site.
+1 hour 30 minutes	The situation is now under control, the people affected are in the process of decontamination and no further risk to the public is perceived at this point. We continue to ask the public to stay away from the area until further notice.
+2 hours 45 minutes	People effected have been decontaminated and are being supported by the team on site. If you are looking for a loved one please contact your local authority through existing channels.
End exercise	The exercise is now over!

Table 1 In-app notifications

Questionnaire

As part of a larger questionnaire of 43 questions entitled the Observer Guide (Carbon et al., 2022), this paper focuses on the questions related to the PROACTIVE mobile app only. The entire Observer Guide was given ethics approval by the PROACTIVE Project Ethics Officer under approval reference: no15/ 26.04.22. Observers were asked to sign an informed consent sheet which explained that their participation in the research was completely voluntary, could be stopped at any time, and that filling in the questionnaire was not mandatory. The questionna ire was distributed in paper format.

The questions on usability and usefulness of the app were developed based on relevant disaster app literature (e.g., Tan et al., 2020) and the available features of the existing app. The questionnaire was composed of 16 questions, with answers on a Likert-type scale ranging from strongly disagree (1) to strongly agree (6) for questions 1 to 12, from not at all useful (1) to very useful (6) for question 13 to 16 on the app features. Respondents also were asked to explain their answer in an open-answer format for all questions.

The first questions focused on aspects of usability:

1. I felt confident using the app
2. The app design is easy-to-use
3. Most people would learn to use the PROACTIVE app quickly
4. The app has effective accessibility features
5. The amount of text displayed was appropriate
6. The visualisations were appropriate
7. It was easy to find critical information about the incident (e.g., time, location, severity)
8. I was able to find information resources/ materials on the topic of CBRNe

The next set of questions focused on usefulness:

9. I would use the PROACTIVE app in the case of a real CBRNe incident
10. I was confident that the incident information I saw on the app was the most recent update
11. The PROACTIVE app enhances the situation awareness of the population on CBRNe events
12. The app respects my privacy (e.g., the privacy statement, GDPR obligations)
13. In-app notifications
14. Incident list
15. Maps showing incidents
16. CBRNe Information Library

The questionnaire responses have undergone a preliminary analysis, both quantitatively and qualitatively. For the Likert scale answers, the average rating is reported, and a higher average score represents a better performance. The content of the open answers provided by the observers were analysed and reported in a qualitative way. The focus was to understand why certain observers provided a lower rating and what were their suggestions for improvement.

Sample Description

The majority of respondents represent CBRNe practitioners (61%) (Figure 2). The majority of the respondents (71%) were also unfamiliar with the PROACTIVE app before the exercise (Figure 3).

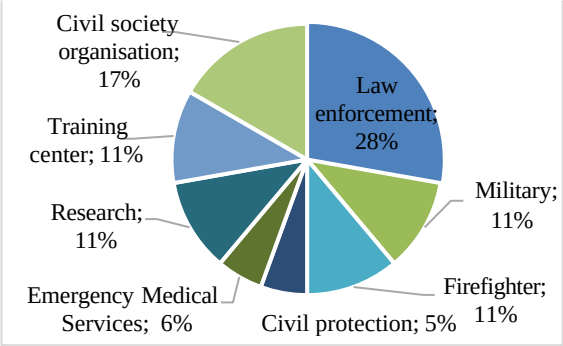


Figure 2. Category of the respondents

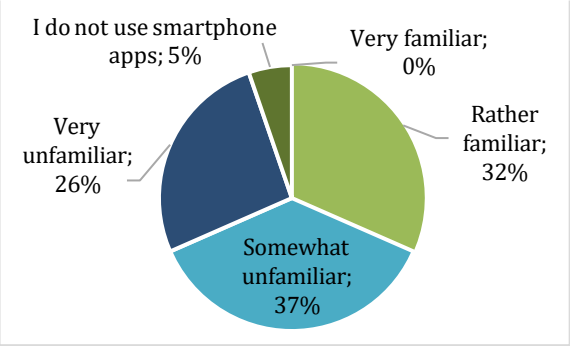


Figure 3. Familiarity with the PROACTIVE app before the field training exercise

RESULTS

Overall, the PROACTIVE app averaged the score of 4.0 on usability (see Table 2 for detailed answers) and 3.9 (see Table 3 for detailed answers) on usefulness.

Question	Number of responses	Average Score	Qualitative Feedback
1	16	3.2	Registration issues, no time to understand the app
2	14	3.9	Simple to use, some technical issues (crashes)
3	14	4.2	Simple to use, concern for the elderly, technical issues (crashes)
4	11	4.0	Suggestions to include language flags and more symbols
5	12	4.3	Well balanced, some translation issues
6	11	4.2	Use of pdf files not appropriate, map to include additional layers
7	13	3.4	Language options to be clearer, use images for incident updates
8	12	4.5	Easy to find, clear and useful information. Use of pdfs advised.

Table 2 PROACTIVE app usability

Question	Number of responses	Average Score	Qualitative Feedback
9	14	3.0	Yes, if linked to national processes and available in local languages
10	13	3.3	Technical Issues prevented this feature being tested
11	13	3.9	Suggestion to be part of existing national processes and apps.
12	12	4.6	Suggestion to include privacy audit and compliance with standards
13	12	3.9	Notifications are useful to verify incidents, but must work flawlessly
14	11	3.7	Useful for multiple stakeholders, suggestion to focus on only live incident and possibility to categorise and/or filter incidents
15	12	4.5	Maps negate language barriers
16	12	4.3	Good feature, only for use outside incident

Table 3 PROACTIVE app usefulness

DISCUSSION

Technical Issues on the Day Appear to Have Impacted Both Usability and Usefulness Scores

On the day of the field training exercise, the PROACTIVE app experienced some (serious) technical difficulties, i.e. the app crashed or froze, clicking on notifications led to an error message, the scroll option was disabled for some users, users who had not yet registered or had forgotten their passwords found it difficult to access the app. However, with support from the on-site technical team, 90% managed to successfully gain access. While not an unusual occurrence for the first time an app or other new technology is tested in a real-life setting, this understandably would influence the questionnaire responses received.

Indeed, the qualitative analysis demonstrates that this is what influenced many of the lower scores, with open-ended answers of "it did not work" or "it kept crashing/freezing" given throughout the questions on usability. Other respondents were able to look past the technical difficulties, for example one observer gave a score of 5 and left an open-ended response of "if it worked" to Q3. These respondents tended to provide rather positive qualitative feedback, such as "the app is self-explanatory," "it is intuitive" or "it is easy to find things." Hence, with an average usability score of 4, it would seem that the observers felt that the PROACTIVE app was usable despite the technical issues, or would be usable if it had worked properly.

These technical issues seem to also have impacted the reported usefulness of the PROACTIVE app. Indeed, if the notifications were not working, and the app kept crashing, it is understandable that observers might not have been sure if the information they saw was from the most recent update (Q10) and were critical of the in-app notification feature (Q13), commenting that it would need to "work flawlessly" to be useful. That said, the average usefulness score of 3.9 seems to indicate that observers felt the app would be, once again, useful despite the technical issues experienced during the field training exercise.

PROACTIVE App Features Could Be Improved

The features were overall deemed useful by participants and it also averaged highly on privacy concerns. With an average of 4.2 for effective accessibility features, it is clear that improvements should be made. Furthermore, qualitative feedback on other questions, not directly related to accessibility, also pinpointed ways forward to ensure the app is developed in line with the principles of universal design, such as using symbols and images to replace the heavy reliance on the written word, avoiding language barriers and making it available in multiple languages. The suggestion to use a country flag to depict which language the app is currently being shown in will be taken forward in the app redesign process.

Usefulness of Yet Another Disaster App Questioned

With a score of 3, it can be said that respondents do not believe they would use the app in the case of a real CBRNe incident. The main qualitative feedback to explain this score was linked to the fact that the app is, as one observer put it, "not connected to national system," and others weren't sure of its added value when compared to other, already existing disaster apps or social media platforms.

This is in line with previous research findings on crisis communication, social media and disaster apps. During a crisis, people tend to turn to the communication channels they use in their everyday life to find information about the crisis (e.g. Steelman et al., 2014), and since disasters are hopefully rare, it can be assumed that few people will have downloaded such an app prior to the given crisis (Petersen et al., 2018). Indeed, a recent representative study of German citizens found that only 16% have downloaded a disaster app (Kaufhold et al., 2020). Furthermore, many studies demonstrate that users prefer to have a single app which would cover all hazards and threats, rather than a single app for each different one (e.g., one app for earthquakes, one app for terrorist attacks) (e.g., Dallo & Marti, 2021; Kaufhold et al, 2020; Haunschild et al. 2022).

As far as being integrated into a national system, the PROACTIVE app is intended to be administered by the local LEA and is being designed so as to easily integrate with legacy systems, thus hopefully meeting this end-user requirement.

FURTHER WORK

In terms of research, it is foreseen to delve deeper into the data presented in this work-in-progress paper, for instance by comparing the responses from CBRNe practitioners to those of civil society organisations.

Furthermore, as part of the iterative, co-creation cycle, the avenues for improvement identified by the questionnaire respondents are planned to be implemented in a new version of the PROACTIVE app. This newer version of the app will then be tested in the second PROACTIVE field training exercise, whereby once aga in feedback will be collected from observers. The third and final PROACTIVE field training exercise will be run using the same methodology. After the last round of improvements, the PROACTIVE app will be considered as the finalised version, concluding the research and development in the context of the PROACTIVE project.

CONCLUSION

This work-in-progress paper presents the preliminary results from a questionnaire on the usability and usefulness of the PROACTIVE CBRNe disaster app, filled in by observers who tested the app during a field training exercise. Overall, observers felt that the app is slightly usable and slightly useful. Their constructive qualitative feedback will be used to improve the app, which will then be tested in an iterative manner in two more field training exercises.

When the project is completed, the PROACTIVE CBRNe app is expected to fill the gap in a lack of citizen oriented CBRNe disaster apps available on the market today. It will be in line with both civil society and CBRNe practitioner's expectations thanks to the co-creation process and should provide the opportunity to minimise the wider "accessibility gap" in emergency management.

ACKNOWLEDGMENTS

We thank all PROACTIVE Dortmund field training exercise observers for their inputs.

Project PROACTIVE has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 832981.

REFERENCES

Bossu, R., Roussel, F., Fallou, L., Landès, M., Steed, R., Mazet-Roux, G., Dupont, A., Frobert, L., & Petersen, L. (2018). LastQuake: From rapid information to global seismic risk reduction. International Journal of Disaster Risk Reduction 28, 32-42. doi: 10.1016/j.ijdrr.2018.02.024.

Carbon, D., Arnold, A., Siemens, M., & Görgen, T. (2021). Common approaches between the vulnerable members of the civil society. Deliverable 3.4 of the PROACTIVE project. https://proactive-h2020.eu/wp-content/uploads/2021/03/PROACTIVE_20210226_D3.4_V5_DHPol_Common -approaches-civil-society.pdf

Carbon, D., Arnold, A., Wüller, C., Görgen, T., Godwin, T., Hale, N., Kelly, D., Marsh, I., Petersen, L., Havarneanu, G., Dennis, A., Carter, H., Weston, D., McCrone, N. & Zamorano, M. Report on the first exercise and evaluation workshop. Delierable 6.3 of the PROACTIVE project. https://proactive-h2020.eu/wp-content/uploads/2022/07/PROACTIVE_20220630_D6.3_V4_DHPol_Dortmund -Field-Exercise.pdf

Connell, B. R., Jones, M., Mace, R., Mueller, J., Mullick, A., Ostroff, E.,Sanford, J., Steinfeld, E., Story,M., & Vanderheiden, G. (1997). The Principals of Universal Design. NC State University, The Center for Universal Design.

Dallo, I. & Marti, M. (2021). Why should I use a multi-hazard app? Assessing the public's information needs and app feature preferences in a participatory process. International Journal of Disaster Risk Reduction 57 (2021), 102197–102197. Issue November 2020. https://doi.org/10.1016/j.ijdrr.2021.102197

European Commission (2019). Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services (Text with EEA relevance), OJ L 151, 7.6.2019.

Gavel, A., Kroupa, T., Navrátilová, L., Setnička, M., Clutterbuck, L., Petersen, L., Havârneanu, G., & Arnold, A. (2021). Recommendations on how to adapt SOPs and tools. *Deliverable 2.4 of the PROACTIVE project*. https://proactive-h2020.eu/wp-content/uploads/2021/03/PROACTIVE_20210228_D2.4_V5_PPI_Recommendations-on-how-to-adapt-SOPs-and-tools.pdf

Gjøsæter, T., Radianti, J., & Chen, W. (2019). Understanding Situational Disabilities and Situational Awareness in Disasters. In Z. Franco, J. J. González, & J. H. Canós (Eds.), Proceedings of the 16th International Conference on Information Systems for Crisis Response And Management. Valencia, Spain: Iscram.

Gjøsæter, J, Radianti, J., & Chen, W. (2021). Universal Design of ICT for Emergency Management from Stakeholders' Perspective - A Systematic Literature Review. Information Systems Frontiers. doi: 10.1007/s10796-020-10084-7.

Hall, C., Weston, D., Long, F., O'Sullivan, F., Amlôt, R., & Carter, H. (2020). Guidelines and recommendations for mitigation and management of CBRNe terrorism. Deliverable D1.3 of the PROACTIVE project. https://proactive-h2020.eu/wp-content/uploads/2021/04/PROACTIVE_20210315_D1.3_V5_PHE_Guidelines-and-Recommendations_revised.pdf

Haunschild, J., Kaufhold, M.-A., Reuter, C. Perceptions and Use of Warning Apps – Did Recent Crises Lead to Changes in Germany? In Mensch und Computer 2022 (MuC '22), September 4–7, 2022, Darmstadt, Germany. ACM, New York, NY, USA, 16 pages. Doi: 10.1145/3543758.3543770

ISO (2013). *Societal security — Guidelines for exercises* (ISO Standard 22398:2013. International Organization for Standardization. https://www.iso.org/obp/ui/#iso:std:iso:22398:ed-1:v1:en.

Kaufhold, M.-A., Haunschild, J., & Reuter, C. (2020). Warning the Public: A Survey on Attitudes, Expectations and Use of Mobile Crisis Apps in Germany. In Proceedings of the 28th European Conference on Information Systems (ECIS), An Online AIS Conference, June 15-17, 2020. https://aisel.aisnet.org/ecis2020_rp/84

Petersen, L, Fallou, L., Reilly, P. & Serafinelli, E. (2020) Public expectations of critical infrastructure operators in times of crisis, Sustainable and Resilient Infrastructure, 5:1-2, 62-77, doi: 10.1080/23789689.2018.1469358

Petersen, L., Havarneanu, G., Markarian, G. & McCrone, N. (2019). CBRNe and vulnerable citizens: co-creating an app for that. Proceedings of the 2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), Paris, France, 2019, pp. 1-4, doi: 10.1109/ICT-DM47966.2019.9032967.

Petersen, L., Havarneanu, G., McCrone, N., Markarian, G. & Kolev, G. (2021). Universal Design & the

PROACTIVE project CBRNe app. In Anouck Adrot, Rob Grace, Kathleen Moore, & Christopher W. Zobel (Eds.), ISCRAM 2021 Conference Proceedings – 18th International Conference on Information Systems for Crisis Response and Management (pp. 959 –966). Blacksburg, VA (USA): Virginia Tech. https://idl.iscram.org/files/laurapetersen/2021/2388_Laura Petersen_etal2021.pdf

Petersen, L., Havarneanu, G., McCrone, N., Markarian, G., Burlin, Å., Johansson, P-E. (2022). CBRNe, a universally designed app for that? In Hedi Karray, Antonio De Nicola, Nada Matta, Hemant Purohit (Eds.), ISCRAM 2022 Conference Proceedings – 19th International Conference on Information Systems for Crisis Response and Management. Tarbes, France.
https://idl.iscram.org/files/laurapetersen/2022/2459_Laura Petersen_etal2022.pdf

Petersen, L., Havarneanu, G., Markarian, G. & McCrone, N. (2023). Practitioner Perspectives of the PROACTIVE CBRNe Disaster App. In J. Radianti, I. Dokas, N. La Lone, D. Khazanchi, eds., ISCRAM 2023 Conference Proceedings – 20th International Conference on Information Systems for Crisis Response and Management (pp. 13 – 19). Omaha, Nebraska, USA: University of Nebraska at Omaha. http://dx.doi.org/10.59297/HVJN6870

Radianti, J., Gjøsæter, J, & Chen, W. (2017). Universal Design of Information Sharing Tools for Disaster Risk Reduction. In: Murayama Y., Velev D., Zlateva P. (eds) Information Technology in Disaster Risk Reduction. ITDRR 2017. IFIP Advances in Information and Communication Technology, vol 516. Springer, Cham. doi: 10.1007/978-3-030-18293-9_8

Steelman, T. A., McCaffrey, S. M., Velez, A. L. K., & Briefel, J. A. (2014). What information do people use, trust, and find useful during a disaster? Evidence from five large wildfires. Natural Hazards, 76(1), 615 –634. doi:10.1007/s11069-014-1512-x

Tan, M. L., Prasanna, R., Stock, K., Doyle, E. E. H., Leonard, G., & Johnston, D. (2020). Understanding end-users' perspectives: Towards developing usability guidelines for disaster apps. Progress in Disaster Science, 7. https://doi.org/10.1016/j.pdisas.2020.100118.



Integrated Wildfire Risk Management

Track Chairs

Claudi Berchtold, Fraunhofer Institute INT

David Martin, Pau Costa Foundation

CoRe Paper – Integrated Wildfire Risk Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Uncertainty in Forest Fire Management

Nikola Tietze
Freie Universität Berlin

Thomas Kox*
Weizenbaum Institute
thomas.kox@weizenbaum-institut.de

Till Büser
Weizenbaum Institute
till.bueser@weizenbaum-institut.de

Lars Gerhold
Technische Universität Braunschweig
lars.gerhold@tu-braunschweig.de

ABSTRACT

In forest fire management, uncertainties are inherent. One major challenge for forest fire managers is the uncertainty of whether a fire will start at all, where, when and with what effect. In particular, human negligence and arson are unforeseeable. This research focuses on these uncertainties affecting the preparatory practices of forest fire managers in Brandenburg, Germany, and how they are anticipated and dealt with. Qualitative content analysis of expert interviews and focus group discussions generate in-depth insights into the role of uncertainties in the prevention phase and during the control of the event. Findings shows that uncertainties pose barriers to preparatory action. Coping mechanisms rely on individual sensemaking that are in turn influenced by one's role, knowledge and experience as well as organizational habits and values. Analysis of German firefighters' perceptions of and coping with uncertainty shows that practices of imagination help to anticipate uncertain futures and to cope with dynamic, complex disasters.

Keywords

embodied uncertainty, forest fire, emergency management

INTRODUCTION

In the management of forest fires, uncertainties are inherent. One major challenge for emergency managers is the uncertainty of whether a fire will start at all, where, when and with what effect (McCaffrey et al., 2020). Forest fire events cannot be predicted with absolute precision, since they arise from a complex of causes consisting of physical and human factors, the actual cause of ignition and the predisposition of the affected areas (Su¨ssner, 2020). Only the potential of the environment to spread the fire, e.g., soil condition, fuel potential and weather facilitating the burning, is used to calculate fire risk such as fire weather forecasts. Thus, uncertainties in fire management previously have mainly been examined in quantitative calculations and models (e.g. Barros-Daza et al., 2022; Diao et al., 2020; Doyle et al., 2019; Wibbenmeyer et al., 2013). These predictions have their limits, as they are subject to forecast uncertainty (Neisser & Kox, 2021; Owen et al., 2012). According to the German forest fire statistics, human action (negligent or intentional arson) is the main cause for forest fires (Bundesanstalt für Landwirtschaft und Erna¨hrung, 2020). This makes the anticipation of such events very challenging. It is therefore important to consider how uncertainties affect the decision-making of forest fire managers and how they cope with these uncertainties (Taylor et al., 2013, p. 7).

Uncertainty is typically split into epistemic uncertainty, which arises from a lack of knowledge, and aleatory or variability uncertainty, which describes unpredictable variables (e.g. Aven, 2011; Kox et al., 2015; Sword-Daniels et al., 2018). Although many models assume that decision-making is predominantly analytic based on calculations and logic, the affective decision-making oftentimes overrules logic if both processes disagree. To enable a more representative concept of decision-making processes, the concept of embodied uncertainty considers the perception

*corresponding author

and handling of uncertainty form a multi-faceted perspective, including emotions and experience (Sword-Daniels et al., 2018). Embodied uncertainty aims to incorporate all aspects of uncertainty in the context of disaster management decision-making and calls for accepting uncertainty as an inherent part of the disaster management task (Sword-Daniels et al., 2018).

Using a qualitative approach, this paper dives deeper into constructions of social reality and thereby reveal underlying belief sets and processes of sensemaking that become relevant when dealing with uncertainty in forest fire management. By doing that, we investigate how uncertainty is embodied. The integration of socio-cultural factors into forest fire management models is still rare. It has previously been attempted by Patra˜o (2020), who calls for a more comprehensive conceptualization of forest fire management that includes socio-cultural realities and a multitude of stakeholders in all phases (e.g. not only firefighters but also the public, the government and agricultural professionals) and emphasizes the role of humans in forest fire management. Through a qualitative perspective this paper aims to explore uncertainties that forest fire managers perceive prior to and during the handling of a forest fire. Furthermore, it reveals how these practitioners contextualize and deal with these uncertainties and which strategies they use (or do not use) to handle the disaster.

Operational decisions are never based on a single piece of information, especially when decision-making is done in the context of multiple factors of uncertainty (Roncoli et al., 2012). Anderson (2010) identifies three practices of anticipation: imagination, performance and calculation. Practices of calculation are limited in forest fire management. Important tools for calculating the hazard risk of forest fires are fire danger rating systems, which are numerical or qualitative indicators that reflect the fire potential (Goldammer, 1998; Neisser & Kox, 2021). These include, in particular, meteorological indices, which combine several influencing factors in one index. In Germany, the forest fire hazard index (WBI) is provided by the agro-meteorological department of the German Weather Service (DWD) indicating daily and weekly forest fire danger levels from 1 (very low danger) to 5 (very high danger) for each county based on air temperature, relative humidity, wind speed and precipitation. Performative practices of anticipation in disaster management are for example training exercises, imitating and therefore embodying possible futures. Regular firefighting drills are part of the training of firefighters, but the focus on specifics of forest firefighting is still rare. Imagination as anticipatory practice refers to envisioning certain futures through scenario planning, and exchanging visions and stories of past events (Anderson, 2010; Neisser & Kox, 2021).

Another concept investigating the perception and management of the unexpected is sensemaking (Weick & Sutcliffe, 2015). In general, sensemaking can be seen as a process of social construction that occurs when discrepant cues interrupt individuals' ongoing activities and involves the retrospective development of plausible meanings that rationalize people's actions (Weick, 1995, 2005). In particular in ambiguous and uncertain moments, sensemaking happens on the account of cues from the environment (Maitlis & Christianson, 2014). Sensemaking processes hence happen in the context of individual and organizational values, experience and knowledge (Barton et al., 2015, p. 81). Whiteman and Cooper (2011) further differentiate that sense-giving while in action plays an important role in managing sudden disasters, such as forest fires. However, unexpected events do not trigger sensemaking per se, but sensemaking occurs when the discrepancy between what is expected and what is experienced is large and important enough to cause individuals or groups to ask "what's going on" and what they should do next. Under growing pressure and prolonged operations sensemaking is deemed to become more challenging and hence, it has been found that under those conditions people rely on their previous understanding of their role and fall back into old habits (Weick & Sutcliffe, 2008).

The purpose of this study is to gain insight into perceptions of uncertainties and barriers to preparedness activities: What uncertainties do forest fire managers perceive in their work? What is their role in forest fire prevention?

MATERIALS AND METHODS

To gain insight into the perceptions of uncertainties and barriers to preparatory measures of forest fire management in Germany, a qualitative research approach was selected. This approach allows to not only shed light on underlying decision-making but also emotions and habits guiding sensemaking and more affective decision-making and reveals in how far uncertainties are embodied. After a systematic literature research, two focus group discussions with overall twelve participants and six in-depth, problem-centered expert interviews were conducted with forest fire protection and management practitioners in Brandenburg, Germany. Brandenburg is the federal state in Germany that is most vulnerable to forest fires, because of its dry soils, same-aged needle woods plantation and continental climate (Gnilke & Sanders, 2021; Thonicke & Cramer, 2006), making it a suitable case study area for assessing forest fire management strategies in Germany. The interviewees included representatives of the Brandenburg Forestry Administration's Forest Fire Monitoring Center and of the Brandenburg State Firefighters Association, employees of regional disaster control administration, as well as local fire chiefs and staff of the control and dispatch centers. The anonymized transcriptions of the interviews were analyzed using qualitative content analysis with

support of MaxQDA software. Qualitative content analysis is particularly useful for gaining detailed insight into the thought processes, meaning-making, and overarching narratives of individuals and groups, and is therefore an appropriate method for uncovering the uncertainties perceived by forest fire managers. Coding was done using a combination of inductive and deductive categories, guided by the relevant literature and taking into account, among other things, the theoretical foundations of organizational sensemaking, organizational culture and management under uncertainty.

RESULTS

The results show that dealing with uncertainty is an important part of the work of forest fire managers. Uncertainty is generally seen as a normal influencing factor of work, which one simply has to accept, since forest fire managers see little possibility to influence it. One can classify the mentioning of uncertainties into uncertainties prior to and during a forest fire event. After examining the general uncertainties that came up during the interviews, we describe coping strategies that have been mentioned or became clear indirectly. The focus lies on the use of forecasting and the role of organizational knowledge. As a final step, we put the results together in a visual overview, attempting to systematize the findings and relate them to the theory discussed above.

Uncertainties Identified Prior to and During the Forest Fire Event

Generally, forest fire managers acknowledge that uncertainties are inherent to forest fire events and mention several uncertainties affecting their work. Overall, they accept that in uncertainty is an invariable inherent component of their work. The uncertainties mentioned by the forest fire managers can be categorized into uncertainty perceived in the phases prior to and during the event (see Tab. 1). Prior to the event uncertainties revolve mainly around the occurrence of a fire and during the event they mostly concern fire behavior (see Fig. 1).

Table 1. Uncertainties prior to and during a forest fire event

Prior to the event	During the event
Weather condition & fuel potential	Availability of volunteer fire fighters
Cause (human ignition)	Development of weather conditions (& fire)
Time and place of the event	Duration of the event

In the anticipation of a possible forest fire, practitioners describe uncertainty mainly as the unpredictability of the time and place of forest fire events, especially since most forest fires in Brandenburg are started by people due to negligent or intentional arson.

Although there is generally an understanding that even weather forecasts cannot predict the upcoming weather conditions in their region with 100% accuracy, forest fire managers still wish for them to be as precise as possibility. Especially the spatial accuracy considering regional conditions and the visualization of meteorological values and their probability are important for practitioners. They state wind direction and speed, wind gusts potential, precipitation, heat, UV radiation and soil moisture as most important variables for them. The ambiguity of trust in weather forecasts becomes clear. On the one hand, the tendency to distrust weather forecasts is due to past deceptions and perceived errors in forecasting; on the other hand, forecasts represent at least one means of computational practice to predict an unexpected event. The actual forecasting of a forest fire event is perceived as highly uncertain and forest fire managers tend to check forecasts through their own observations or through information from people on site.

After the alarm to a forest fire, the biggest uncertainty lies in the structure inherent to the fire management system, namely its dependency on volunteer firefighters. Because of the sudden alarm, it is difficult to know exactly how many of the firefighters can arrive in time at the forest fire event. A major problem with this is that many volunteer firefighters today work in larger cities that are far from their home. This limits their ability to arrive at their assigned fire station or on site in the shortest time possible. Another unknown variable is the duration of the forest fire, as it depends heavily on the time of arrival of the first firefighters, weather conditions such as wind speed and direction, soil moisture and precipitation, and the local conditions of the affected area. In the region, there is also the problem that some forests are former military sites and therefore there are restrictions on entering the area to ensure the safety of firefighters.

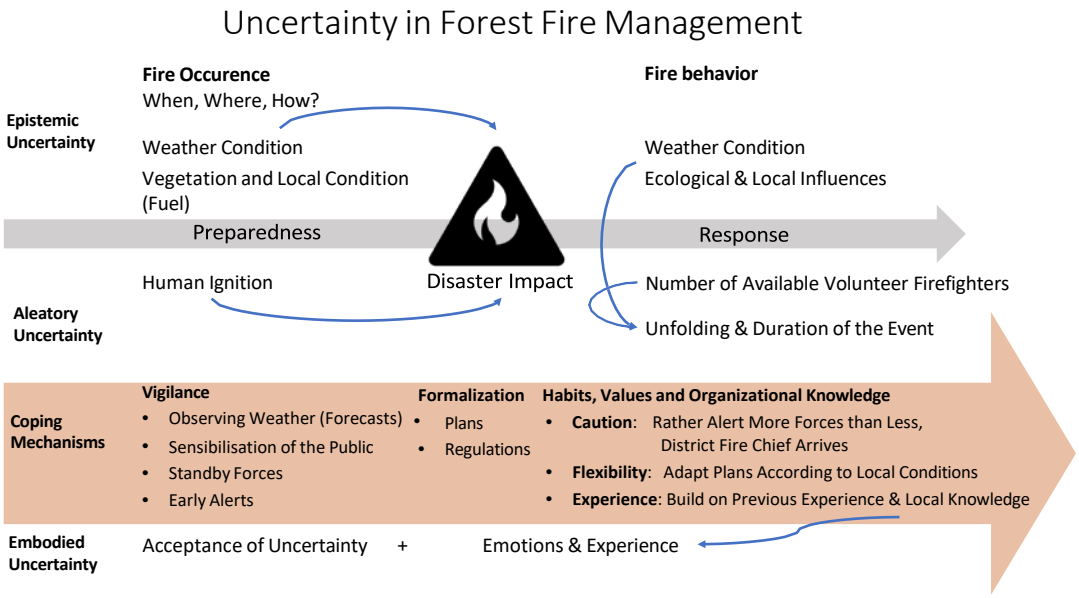


Figure 1. Uncertainty in forest fire management and coping mechanisms to deal with these uncertainties

Coping Mechanisms for Dealing with Uncertainties

The interviews revealed different strategies for coping with uncertainty. While tools and probabilistic calculations such as weather forecasts and formalized structures such as plans and established procedures minimize epistemic uncertainty, values and experiences as part of organizational culture seem to play an important role in managing aleatory uncertainty.

Vigilance

To minimize the impact of uncertainties on forest fire management, most forest fire managers seem to be increasingly vigilant towards cues for forest fire risks. Most importantly, weather forecasts and forest fire risk indices are a relevant tool to envision possible origins of forest fires and even more so the potential development of the forest fire behavior. Forest fire managers rely mostly on calculations such as meteorological weather information because regulations and legal liability hold them responsible for actions taken or not taken based on different forecasts or fire warning levels of the German forest fire risk index. Probabilistic forecasts are sometimes mentioned to help consider those uncertainties in an even more detailed way. Nevertheless, the aim to counter the uncertainties of weather forecasts by predicting possible impacts through impact-based forecasting is oftentimes considered irrelevant by the interviewees, who emphasize that experience and local knowledge are their most important guidance. However, forest fire managers see potential in providing such additional guidance to the public, who they believe are not as well informed. The interviewees deem that the public could be more inclined to seek protection and change their behavior when receiving additional information on possible weather impacts. This in turn would ease their job, as fewer operations would become necessary if people were more sensitized to the risk of fire weather and acted in accordance with the warnings. Another aspect of vigilance and increased alertness is to call standby forces to the fire station, if the fire risk is assessed to be high according to the weather forecast, the forest fire warning index, or personal feeling. However, as there are no formal regulations to allow this standby procedure and most German firefighters are volunteers, this is only possible on weekends or evenings, when the firefighters do not have to attend their regular job. Otherwise, firefighters rely on early alerts to get to the scene as quickly as possible.

Formalization

Another way to cope with uncertainties is to prepare for events through plans based on previous experiences and risk calculations. In Brandenburg, Germany, there is a defined organizational structure specifying who is in charge of making plans, revising and enacting these plans. These regulations and guidelines are distributed in the emergency vehicles, the command and dispatch center and at the fire stations. Commonly, plans are to be updated every three years or after extreme forest fires. There are also forestry regulations that allow forest managers to restrict public access to the forest when there is a high forest fire warning level of 4 or 5 out of 5. However, this is not commonly used by the practitioners, because the forest managers feel that these regulations cannot be enforced.

Habits, Values and Organizational Knowledge

To balance uncertainties, some general organizational and individual habits come into play. One of them is increased caution. According to the interviewed fire managers, the district (or county) fire chief (*Kreisbrandmeister*) always takes part in the first forest fire assessment onsite if the notifying command and dispatch center alerts a bigger forest fire with the code for larger forest fires (*Brand: Wald/groß*). The motto is to rather request more emergency personnel than necessary than have too few available. Since one of the greatest uncertainties is that the volunteer firefighters will actually show up in time. Because of that, some fire brigades even choose to call a stationary attendance if they suspect an event happening in the close future, although no formal plans to call such stand-by attendance exist. Another aspect of taking caution is to forward every warning received, even if it is a false alarm. Caution during the incident particularly involves keeping the firefighters safe, e.g. from threats arising from old munition, thunderstorm, heat and high UV or landslide risk. Especially during a bigger event some interviewees deem it relevant to stay calm. For example, one interviewee emphasizes that he learned from older colleagues to keep calm and wait, and manage a the event with composure. Other practitioners emphasize that there is "no universal dogma" applicable to any situation and all fire stations as "no operation is the same as another". This demonstrates that adaptivity and flexibility are relevant to coping with forest fires, their unexpected cause and their uncertain development. A positive attitude and motivation for their job support these habits and values. All practitioners agree that experience is more important than any predictive calculation, weather forecast, or forest fire risk index. As one fire brigade leader puts it, firefighting is always a "step by step" learning process, in which all firefighters evolve and enhance their skills. This implies that knowledge transfer within the organization is of high relevancy. However, knowledge transfer is rarely systematic. In some organizations, certain stories are told by senior colleagues, while in others it is emphasized that experience is best gained by being exposed to real firefighting operations, not just practical exercises or theoretical lessons. Talking to other members of the organization and learning from their experience is thought to be more helpful than learning by book. Some use a mentoring program that gradually gives younger, newer firefighters more responsibility in the field, but always offers the chance to draw on the experience of the mentor in case of difficulties. At the governance level, knowledge exchange also extends to international and scientific exchange. Successful firefighting is also based on mutual trust. Local fire brigade leaders create this trust by repeatedly making it clear that the safety of the firefighters is paramount, by promoting an open culture of error and being proud of their comrades.

Integrative Framework of Uncertainty Perceptions and Coping Mechanisms

Fig. 1 visualizes which types of uncertainties forest fire managers perceive, how they can be theoretically categorized and which organizational coping strategies are employed to encounter them. The uncertainties and their coping strategies are organized along their stage of appearance in the disaster cycle. Interactions and dependencies are presented as blue arrows. For example, the uncertainties of the fire occurrence also affect the disaster impact. Furthermore, uncertainties about the fire behavior that we categorized as epistemic such as the weather condition predicted by forecasts, but also aleatory uncertainties such as the number of volunteer firefighters actually showing up, influence another factor of uncertainty: the unfolding and duration of the event. The coping mechanisms primarily applied in the different phases can be categorized into the anticipatory practices, which help deal with the different types of uncertainties and at different disaster stages (see 1). At different stages of the disaster cycle different practices of anticipations have varying relevancy to anticipate uncertainties and cope with the uncertainties. Furthermore, the type of uncertainty seems to influence the type of anticipation and coping strategy .

DISCUSSION

In the literature on forest fire management, it has been widely acknowledged that weather is an important factor in the development and modification of forest fire events (Boychuk et al., 2020). This study goes further and assesses the perception of weather forecast uncertainty in context with further uncertainties affecting forest fire management decision-making. Weather forecasts and warnings are important for practitioners, especially because of the related legal regulations. For the actual firefighting, however, they are combined with personal experience, own observations and local knowledge in order to make optimal adapted decisions. Hence, when distinguishing between intrinsic and contextual factors of operational decision-making (Roncoli et al., 2012), particularly the contextual factors of decision-making seem to be of relevance in forest fire management, as the situation is very dynamic and decisions have to be assessed and adapted accordingly.

When uncertainties prevail, the decision environment in which predictions are used matters. The results of our research show that apart from formalized structures in forest fire management, such as the organizational structure, information flows, or human factors especially values, previous experiences and local and organizational knowledge affect decision-making processes under uncertainty in forest fire management. To respond to the uncertainty

inherent to their job, organizations tend to follow habits, plans, goals or rituals that are intended to create more certainty and control over the situation, well-knowing that they can never be sure of the success of these aims (Turner, 1976, p. 378). The relevancy of values, experience and knowledge in sensemaking processes has also been demonstrated to be crucial by former studies (Barton et al., 2015, p. 81). In our case study, especially experience and local knowledge were significant factors to cope with uncertainties in forest fire management. This finding agrees with Weick and Sutcliffe's 2008 analysis of forest fire managers' sensemaking practices, who state that under growing pressure sensemaking becomes more challenging and hence people rely on their previous experiences and understanding of their role.

The analysis shows that values such as flexibility and adaptivity and a calm mindset are deemed crucial to cope with uncertainties in forest fire management. This finding can be used to develop these skills in forest fire exercises. Especially the aspect of flexibility and openness to adapt the current organizational structures and decision-making to the dynamic situation has been considered beneficial for emergency management by previous scholars and has been highlighted especially for firefighting (Weick & Sutcliffe, 2008).

Another mechanism that German forest fire practitioners emphasized for coping with uncertainty is increased vigilance and caution. Practitioners remain attentive to indications that the situation may change and not unnecessarily jeopardizing the safety of responders. Firefighters in Germany deem it impossible to completely hinder the occurrence of an event, but they always aim to get ahead of the threat and change the situation (Kox & Lüder, 2021). The forest fire managers tendency to rather warn and alert too often than too little has for instance also been observed in a study on hurricane prevention management, calling this practice the "worst-cased approach" (Bostrom, A., 2016), or in firefighters' response to weather warnings, calling it a "zero-risk strategy, in which even the lowest probability of occurrence is too much of a risk and thus has to be avoided" (Kox et al., 2018, p. 126). The safety of responders is a key priority in managing the event and previous studies have recognized the influence of self-protection decisions on fire suppression strategies (e.g. Patra˜o, 2020). The interviewees in this study also mentioned the aim to reach out to the public as part of their preparatory measures trying to prevent forest fires from happening. It seems that practitioners in fact already include socio-cultural factors in their management strategies and acknowledge the value of incorporating local communities into forest fire management (Patra˜o, 2020; Paveglio et al., 2019).

Barton et al. (2015, p. 81 find as an additional factor to manage uncertainty that leadership behavior influences effective disaster response. This aspect has not come up much with the interviewees in this context. One reason why this was only rarely mentioned could be that all interviewed persons were in a leadership position, and hence did not reflect the importance of their position and influence. Another reason could be that organizational structures and culture, and experience just matter more for forest fire managers than leadership behavior. However, the research shows that knowledge transfer is another important factor for managing uncertain complex fire events and the effectiveness of organizational knowledge transfer probably relies on leadership. The unclarity of the role of leadership in this research opens room for further research that should investigate on the links of knowledge transfer as part of leadership behavior, and leadership's potential to minimize uncertainties. Previous research has already discussed the impact of leadership on firefighters well-being but not focused on perceptions of uncertainty (Marques-Quinteiro et al., 2022).

Overall, our research has shown that to cope with uncertainties, practices of imagination, such as knowledge transfer, experience of senior colleagues told through stories or organizational values shaping individual sensemaking are highly relevant. These habits and values are shaped by a general acceptance of uncertainty. This finding agrees with previous research that asserts professional weather forecast users relatively good knowledge of forecast uncertainty (Kox et al., 2015). This embodied uncertainty helps to face the emerging threats with composure and to remain calm and flexible even in extreme situations. Accordingly, the interviewees stressed the relevancy of personal experiences, mindsets and emotions. As this is an exploratory case study, the research is limited to providing deep insights into the expert group in question; it cannot be expected that practitioners in other fields or other cultural contexts will necessarily hold the same views. Moreover, qualitative research is never completely free of subjectivity. Employing Lincoln and Guba's Standards for qualitative research in respect to transparency, credibility, dependability and confirmability was a measure to overcome this challenge (Shenton, 2004). Our results represent the viewpoints of the interviewees, which may be affected by social desirability and the general interview setting but are contextualized in the discussion and interpreted in the light of previous studies and findings of scholars exploring the management of the unexpected.

In contrast to previous quantitative studies, our research sheds deep insights into perceptions of uncertainties in forest fire management from the perspective of practitioners and analyses these in the context of the current state of organizational and disaster management research approaches. The findings serve to reveal underlying reasonings that influence decision-making under uncertainty in forest fire management as well as coping strategies. The

study thereby confirms findings of previous studies by analyzing first-hand experiences and conceptualizations of emergency service personnel in the forest fire sector. Moreover, we showed performance practices such as exercises to be less relevant to the group and highlighted aspects such as the role of local knowledge and knowledge transfer, that so far have received less attention by disaster scientists in Western countries. Presenting these findings will not only help advance research on management strategies in disaster management but also aid practitioners to contextualize and adapt their strategies. Our work adds to the growing literature on anticipatory wildfire governance a qualitative perspectives on use and perception of uncertainties (Miller et al., 2022). Future studies should take these findings and compare them in a wider, more international scope to find differences and similarities. This will enrich future strategies for managing and accepting the unexpected and integrating embodied uncertainty into forest fire management approaches.

CONCLUSIONS

Our study demonstrated perceptions of uncertainty in forest fire management prior to and during the event. The sensemaking approach was a valuable approach to classify these perceptions and how to deal with them. We show how German firefighters consciously and unconsciously employ certain strategies to deal with uncertainty, e.g., adaptive informal awareness mechanisms or formalization such as making plans based on past events. In general, a wait-and-see caution and constant vigilance are evident. In particularly extreme events, where aleatory uncertainties dominate, computational and performance practices are less relevant to anticipating the future development of a fire, while practices of imagination are essential. Experience, caution, composure, and flexibility are values that have been shown to be relevant to forest fire management in the face of uncertainty. In addition, the role of local knowledge and knowledge transfer in managing unexpected extreme fire events was emphasized and should be further explored in future studies.

ACKNOWLEDGMENTS

This research was carried out in the Hans-Ertel-Centre for Weather Research. This network of Universities, Research Institutes and the Deutscher Wetterdienst is funded by the German Federal Ministry of Transport and Digital Infrastructures (DWD2014P3A). The contribution of Thomas Kox and Till Bu¨ ser was part of the joint project Weizenbaum Institute for the Networked Society (16DII131), funded by the German Federal Ministry of Education and Research.

REFERENCES

Anderson, B. (2010). Preemption, precaution, preparedness: Anticipatory action and future geographies. *Progress in Human Geography*, *34*, 777–798. https://doi.org/https://doi.org/10.1177/0309132510362600

Aven, T. (2011). On Different Types of Uncertainties in the Context of the Precautionary Principle: On Different Types of Uncertainties in the Context of the Precautionary Principle. *Risk Analysis*, *31*(10), 1515–1525. https://doi.org/10.1111/j.1539-6924.2011.01612.x

Barros-Daza, M. J., Luxbacher, K. D., Lattimer, B. Y., & Hodges, J. L. (2022). Real Time Mine Fire Classification to Support Firefighter Decision Making. *Fire Technology*, *58*(3), 1545–1578. https://doi.org/10.1007/s10694-022-01215-4

Barton, M. A., Sutcliffe, K. M., Vogus, T. J., & DeWitt, T. (2015). Performing Under Uncertainty: Contextualized Engagement in Wildland Firefighting: Performing Under Uncertainty: Contextualized Engagement in Wildland Firefighting. *Journal of Contingencies and Crisis Management*, *23*(2), 74–83. https://doi.org/10.1111/1468-5973.12076

Bostrom, A. (2016). A Mental Models Study of Hurricane Forecast and Warning Production, Communication, and Decision-Making. *Weather, Climate and Society*, *8*. https://doi.org/10.1175/WCAS-D-15-0033.s1.

Boychuk, D., McFayden, C. B., Evens, J., Shields, J., Stacey, A., Woolford, D. G., Wotton, M., Johnston, D., Leonard, D., & McLarty, D. (2020). Assembling and Customizing Multiple Fire Weather Forecasts for Burn Probability and Other Fire Management Applications in Ontario, Canada. *Fire*, *3*(2), 16. https://doi.org/10.3390/fire3020016

Bundesanstalt fu¨r Landwirtschaft und Erna¨hrung (Ed.). (2020). Waldbrandstatistik der Bundesrepublik Deutschland fu¨r das Jahr 2019.

Diao, T., Singla, S., Mukhopadhyay, A., Eldawy, A., Shachter, R., & Kochenderfer, M. (2020). Uncertainty Aware Wildfire Management. https://doi.org/http://arxiv.org/abs/2010.07915

Doyle, E. E., Johnston, D. M., Smith, R., & Paton, D. (2019). Communicating model uncertainty for natural hazards: A qualitative systematic thematic review. *International Journal of Disaster Risk Reduction*, *33*, 449–476. https://doi.org/10.1016/j.ijdrr.2018.10.023

Gnilke, A., & Sanders, T. (2021). Forest fire history in Germany (2001-2020). *Eberswalde: Thu¨nen Institute of Forest Ecosystems*, *2*.

Goldammer, J. (1998). Early Warning Systems for the prediction of an appropriate response to wildfires and related environmental hazards. In J. Goldammer (Ed.), *Health Guidelines for Vegetation Fire Events*.

Kox, T., Gerhold, L., & Ulbrich, U. (2015). Perception and use of uncertainty in severe weather warnings by emergency services in Germany. *Atmospheric Research*, *158-159*, 292–301. https://doi.org/10.1016/j.atmosres.2014.02.024

Kox, T., & Lu¨der, C. (2021). Impacts as Triggers for Weather-Related Decision Making: Observations at the Berlin Fire Brigade Control and Dispatch Center. *International Journal of Disaster Risk Science*, *12*(4), 610–615. https://doi.org/10.1007/s13753-021-00356-4

Kox, T., Lu¨ der, C., & Gerhold, L. (2018). Anticipation and response: Emergency services in severe weather situations in Germany. *International Journal of Disaster Risk Science*, *9*, 116–128.

Maitlis, S., & Christianson, M. (2014). Sensemaking in Organizations: Taking Stock and Moving Forward. *Academy of Management Annals*, *8*(1), 57–125. https://doi.org/10.5465/19416520.2014.873177

Marques-Quinteiro, P., Chambel, M., & Maio, A. (2022). Leadership at the Extreme: A Longitudinal Study of Transformational Leadership Style and Well-Being in Firefighters. *Fire*, *5*(6), 192. https://doi.org/10.3390/fire5060192

McCaffrey, S., McGee, T., Coughlan, M., & Tedim, F. (2020). Understanding wildfire mitigation and preparedness in the context of extreme wildfires and disasters. In F. Tedim, V. Leone, & T. McGee (Eds.), *Extreme Wildfire Events and Disasters: Root Causes and New Management Strategies* (pp. 155–174). Elsevier.

Miller, B. A., Yung, L., Wyborn, C., Essen, M., Gray, B., & Williams, D. R. (2022). Re-Envisioning Wildland Fire Governance: Addressing the Transboundary, Uncertain, and Contested Aspects of Wildfire. *Fire*, *5*(2), 49. https://doi.org/10.3390/fire5020049

Neisser, F., & Kox, T. (2021). Zur Verankerung von Zukunft in der Gefahrenabwehr. *Zeitschrift fu¨r Zukunftsforschung*, *1*, 158–181.

Owen, G., McLeod, J. D., Kolden, C. A., Ferguson, D. B., & Brown, T. J. (2012). Wildfire Management and Forecasting Fire Potential: The Roles of Climate Information and Social Networks in the Southwest United States. *Weather, Climate, and Society*, *4*(2), 90–102. https://doi.org/10.1175/WCAS-D-11-00038.1

Patra˜o, A. (2020). *Human and Social Dimensions of Wildland Fire Management and Forest Protection* (L. Land, W. Leal Filho, A. Azul, L. Brandli, A. Lange Salvia, & T. Wall, Eds.). Springer International Publishing.

Paveglio, T. B., Edgeley, C. M., Carroll, M., Billings, M., & Stasiewicz, A. M. (2019). Exploring the Influence of Local Social Context on Strategies for Achieving Fire Adapted Communities. *Fire*, *2*(2), 26. https://doi.org/10.3390/fire2020026

Roncoli, C., Breuer, N., Zierden, D., Fraisse, C., Broad, K., & Hoogenboom, G. (2012). The art of the science: Climate forecasts for wildfire management in the southeastern United States. *Climatic Change*, *113*(3-4), 1113–1121. https://doi.org/10.1007/s10584-012-0526-1

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63–75. https://doi.org/10.3233/EFI-2004-22201

Su¨ssner, B. (2020). Wald - und Vegetationsbra¨nde: Pra¨vention, Einsatzvorbereitung.

Sword-Daniels, V., Eriksen, C., Hudson-Doyle, E. E., Alaniz, R., Adler, C., Schenk, T., & Vallance, S. (2018). Embodied uncertainty: Living with complexity and natural hazards. *Journal of Risk Research*, *21*(3), 290–307. https://doi.org/10.1080/13669877.2016.1200659

Taylor, S. W., Woolford, D. G., Dean, C. B., & Martell, D. L. (2013). Wildfire Prediction to Inform Fire Management: Statistical Science Challenges. *Statistical Science*, *28*(4). https://doi.org/10.1214/13-STS451

Thonicke, K., & Cramer, W. (2006). Long-term Trends in Vegetation Dynamics and Forest Fires in Brandenburg (Germany) Under a Changing Climate. *Natural Hazards*, *38*(1-2), 283–300. https://doi.org/10.1007/s11069-005-8639-8

Turner, B. A. (1976). The Organizational and Interorganizational Development of Disasters. *Administrative Science Quarterly*, *21*(3), 378. https://doi.org/10.2307/2391850

Weick, K. E. (1995). *Sensemaking in organizations*. Sage Publications.

Weick, K. E. (2005). Managing the Unexpected - Complexity as Distributed Sensemaking. In R. McDaniel & D. Driebe (Eds.), *Uncertainty and Surprise in Complex Systems. Understanding Complex Systems* (pp. 51–65). Springer.

Weick, K. E., & Sutcliffe, K. M. (2008). Organizing for Higher Reliability: Lessons Learned from Wildland Firefighters. *Fire Management Today*, *68*, 14–19.

Weick, K. E., & Sutcliffe, K. M. (2015). *Managing the unexpected: Sustained performance in a complex world* (Third edition). Wiley.

Whiteman, G., & Cooper, W. H. (2011). Ecological Sensemaking. *Academy of Management Journal*, *54*(5), 889–911. https://doi.org/10.5465/amj.2008.0843

Wibbenmeyer, M. J., Hand, M. S., Calkin, D. E., Venn, T. J., & Thompson, M. P. (2013). Risk Preferences in Strategic Wildfire Decision Making: A Choice Experiment with U.S. Wildfire Managers: Risk Preferences in Strategic Wildfire Decision Making. *Risk Analysis*, *33*(6), 1021–1037. https://doi.org/10.1111/j.1539-6924.2012.01894.x

Networks for Integrated Wildfire Management: A Social Network Analysis Approach

Diogo Vallim
Copenhagen Business School
dv.msc@cbs.dk

Amalia Giannakopoulou
Copenhagen Business School
agi.msc@cbs.dk

Anna Schmid
Copenhagen Business School
ajcs.msc@cbs.dk

David Struthers
Copenhagen Business School
dst.msc@cbs.dk

Efthymios Altsitsiadis
Copenhagen Business School
ea.msc@cbs.dk

Isabel Fróes
Copenhagen Business School
ifr.msc@cbs.dk

ABSTRACT

Wildfires, exacerbated by climate extremes, pose escalating global threats. Despite advancements in technology, current wildfire management policies lack specific details about integrated management and stakeholder engagement. This study provides empirical evidence on how five European wildfire management locations align with networked governance models by analyzing how their stakeholder networks are configured through social network analysis (SNA). This is measured by assessing network configuration and stakeholder group profiles. Measurements at the network level assess the size, density, diversity of sectors, scalar levels of operation, and phases. At stakeholder level, relevance is assessed by measuringin degree, eccentricity, closeness centrality, and betweenness centrality. Analysis qualitatively assesses the network profile, leading to the discussion on the primary factors influencing network configurations and the potential outcomes of these attributes.

KEYWORDS

Integrated Wildfire Management, Social Network Analysis, Stakeholders Analysis

FUNDING DECLARATION

This research has received funding from the European Union's Horizon 2020 programme under grant agreement No 101036926 (A Holistic Fire Management Ecosystem for Prevention, Detection and Restoration of Environmental Disasters [TREEADS]). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

INTRODUCTION

The frequency of wildfires and the risk they pose is increasing, exacerbated by extreme climate events such as heatwaves and droughts. In the past five years, Portugal (2017), Greece, (2023), and other locations experienced devastating fires, highlighting the severity of the issue. Inadequate disaster and forest management increase the short-term impact and long-term repercussions of wildfires (Marie & Labossière, 2015; Pearce, 2003; Pereira et al., 2016). These events prompted extensive discussion on the necessity of improving existing mechanisms, collaborations, and management approaches. Beyond advancements in technology, stakeholders emphasize the importance of integrated strategies spanning wildfire management phases, focusing on actors, governance, technology, and socio-economic factors. Within the European Union, the integrated approach is promoted through projects funded by Horizon 2020 Innovation Action grants and Union Civil Protection Calls.

Wildfire management is as a collective action problem requiring strong coordination among stakeholders (Charnley et al., 2020). Engaging diverse stakeholders in collaborative processes that integrate environmental information into decision-making is crucial, but also challenging. It is necessary to work at and across the boundaries of knowledge types, which involves navigating complex environments and multiple stakeholder-engagement processes that can promote knowledge exchange and co -creation (Sitas et al., 2016). Despite developments in comprehensive wildfire management, prevailing strategies lack detail about how to identify relevant stakeholders and the purpose of stakeholder engagement. Crucially, there is also little research evaluating existing levels of stakeholder engagement in the contexts where these processes are situated. A changing environment, the need to disseminate information, complex decision-making processes, and insufficient understanding of all actors increase the risk of managerial mistakes (Battistoni et al, 2020).

A central tenant in comprehensive wildfire management involves transitioning from a paradigm focused solely on enhancing the efficiency of civil protection services in response to wildfires to a model that prioritizes risk identification and assessment, coupled with proactive prevention measures. The use of SNA methods in wildfire management research is justified by recent calls from academics, land managers, and policymakers for an approach to wildfire risk management that works across sectors to consider the needs of different stakeholders in risk assessment and communication efforts (AGIF, 2023; Casartelli, & Mysiak, 2023; Chuvieco et al., 2023; United Nations, 2015). These calls to bridge gaps across sectors and scalar levels of operations fits into the debate on integrated wildfire management that began in the1980s. The issue has become more pressing in recent years as the frequency of catastrophic fires increases along with demand that land management processes align with economic, public, and environmental pressures (Ferrara et al., 2019; Moore, 2019).

Successfully implementing this new approach to wildfire risk management is contingent on stakeholder engagement. For example, the significance and effectiveness of governing institutions may not align with their formal roles, or specific information on practices may be held by stakeholders not formally participating in governance. Social network analysis (SNA) is a valuable tool for examining the dynamics of social interactions, which cannot be fully grasped by only examining formal institutions in a natural resource governance system. It provides insights on the most relevant actors within a given a network, how these actors collaborate, and how information flows in systems across different scalar levels in wildfire management contexts within the European Union.

LITERATURE REVIEW

Changes in wildland-urban interfaces, climate change, and inadequate forest and disaster management increase the severity of wildfires and the complexity of wildfire management (Marie & Labossière, 2015; Pearce, 2003; Pereira et al., 2016). A new wildfire management approach is needed to effectively adapt to these changes.

Governments, policy experts, and researchers have extensively discussed the necessity of improving existing governance mechanisms, collaborations, and management approaches to wildfires. Beyond the adoption of new technologies, they emphasize the importance of integrated strategies encompassing all wildfire management phases by focusing on stakeholders, governance, technologies, and socio-economic factors. Understanding wildfire management as a collective action problem that requires strong coordination amo ng stakeholders is key to successfully shifting approaches (Charnley et al., 2020). Engaging diverse stakeholders in collaborative processes involves navigating complex environments (value systems, social conventions, and mental models) through multiple stakeholder-engagement processes that can promote knowledge exchange and co -creation (Sitas et al., 2016). The long-term complexity of these issues is evidenced by what researchers refer to as the "wildfire paradox" created by fire suppression strategies, including public awareness campaigns aimed at preventing wildfires. Fewer fires in the short term raises the fire risk from stored fuel (Calkin et al., 2015).

Limiting management fires also contributes to excess fuel storage. Scholars and policy makers in wildfire-prone regions argued that past approaches to fire suppression and climate change necessitate society creating "a new relationship with fire" (Thompson et al., 2018, p. 382).

Efforts to formulate pathways to change entail realigning priorities and risk sharing among multiple public and private stakeholders and viewing wildfire governance as a "social-ecological problem" (Steelman, 2016). Making these shifts, however, opens complexity that needs to be managed by policy makers, practitioners, and researchers. Relationships between stakeholders and distinct stakeholder understandings of the risks posed by wildfires need to be factored into effective wildfire governance strategies. However, Kirschner et al. (2023) noted imprecise definitions and the lack of a consistent approach to studying wildfire governance. The authors proposed a four-part analytic framework of participation, collaboration and co-production, path-dependencies and place-based dynamics, and adaptation and anticipation.

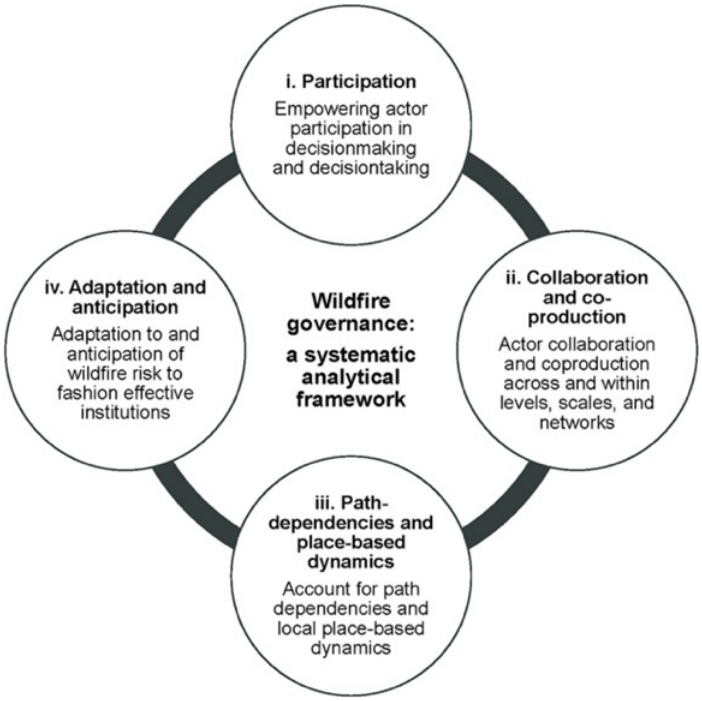


Figure 1: Wildfire Governance Framework (Source: Kirschner et al., 2023)

Kirschner et al. (2023) is an important reframing that provides analytical categories for understanding relationships among stakeholders and governance processes. Research by networked governance scholars resonate with these categories. Provan and Kenis (2007) documented three types of network governance—participant-governed networks, lead organization-governed networks, and network administrative organization—and developed criteria for evaluating the effectiveness of network governance. Addressing complex issues demanding multilateral coordination in areas such as disaster management, economic development, or social services requires achieving the goals of multiple organizations (Provan and Kenis, 2007; O'Toole, 1997).

Bodin & Nohrstedt (2016) built on this knowledge to develop a model to "empirically investigate relationships between task interdependency and collaborative (social) network structures" (p. 184). The authors argued for "the importance of moving beyond simplified assertions of the utility of collaborative governance in addressing disasters by enabling more precise and theoretically informed empirical inquiries regarding the mechanisms that shape the structure and performance of collaborative networks" (p. 192). Carboni et al.'s (2019) examination of purpose-oriented networks draws our attention back to the goal of the project pilots, namely, improving wildfire response, which their work provides theoretical insight to evaluate.

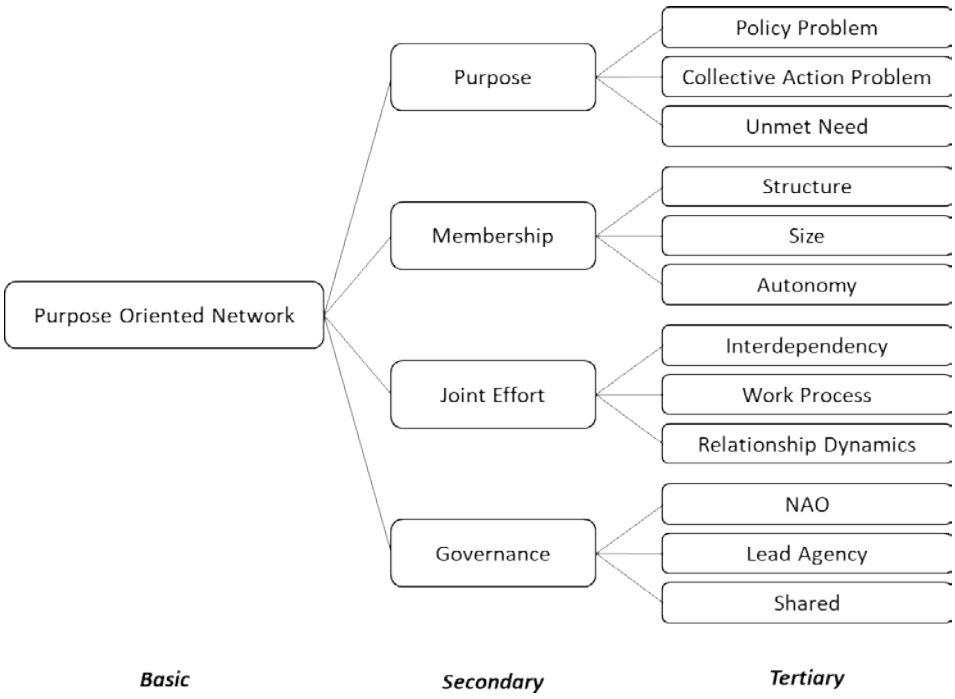


Figure 2: Purpose-Oriented Network Conceptual Framework (Source: Carboni et al., 2019, p. 212)

Charnley et al. (2020) framed wildfires as a "collective action problem in multi-ownership landscapes" (p. 2). Steelman et al. (2021) begin their article by centering the networked aspects of disasters before introducing structuring factors of networked governance during wildfires. De Hoop et al. (2022) contributed to understandings of the multi-stakeholder aspects of wildfire management. Bradshaw (2019) provided guidance through their evaluation of stakeholder-agency interactions in wildfire management (see also Burns et al., 2003). Essen et al. (2023) specifically examine risk sharing in this domain (see also Fischer & Jasny, 2017). Hamilton et al. (2023) contributed findings about risk, wildfires, and multilevel governance.

There are significant policy efforts to implement these networked approaches to wildfire management. The EU's Wildfire Peer Review Assessment Framework (WPRAF) (Casartelli & Mysiak, 2023) and the 8th International Fire Land Conference's Landscape Fire Governance Framework (LFGF) (AGIF, 2023) are two recently formulated governance frameworks created to account for the shifts discussed above. These two holistic frameworks structure governance approaches across wildfire phases, eco -socio-economic factors, and stakeholders. They promote novel approaches to risk governance that changes how information about risk is collected, analyzed, and communicated with the goal of improving risk management decision making. These frameworks position decision making as a negotiated process rather than determined by governmental authorities and break silos, increase collaboration, and prioritize training and preparation. Our study draws from the research discussed above and these two governance frameworks to stress the importance of understanding diverse stakeholders and their relationships in holistic wildfire management.

METHODS

The project partners conducted in-depth stakeholder analysis in support of the project's proposed outcomes of improving wildfire management through the project pilots. Stakeholder mapping and analysis is an effective starting point for identifying stakeholders, understanding their rela tionships, and their positions in holistic wildfire governance approaches. It also supports continuous stakeholder engagement through the project pilots. We approached stakeholder analysis through Reed et al.'s (2009) survey of stakeholder literature in business management and natural resource management cataloging stakeholder analysis methods. Our methodological approach to understanding the multi-actor complexities of wildfire governance treated the pilots as focus groups to work with in order to map stakeholders within and outside of the project pilots. After completing the stakeholder map for each pilot area, we analyzed stakeholder relationships through social network analysis. Reed et al. (2009) characterized the strength of SNA as helping "Gain insight into the boundary of stakeholder network; the structure of the network; identifies influential stakeholders and peripheral stakeholders" (p. 1937).

The stakeholder analysis method employed in this study aims to identify groups and organizations that impact

or are impacted by wildfire regimes in the target regions, develop knowledge about their relationships, and determine their relevance in the wildfire management process. The approach follows Wasserman and Faust (1994) and Carrington et al. (2009) that advocate for the importance of networks as the primary units of analysis. The relationships between stakeholders are of primary interest and the attributes of stakeholders arise from the relational processes in which they are integrated.

Our approach follows Reed et al. (2009) and Ahmadi (2019) work on stakeholder analysis in natural resource governance. Stakeholder analysis is a process that starts with identif ying stakeholders (I), before categorizing (II) and mapping (III) their relationships through social network analysis (SNA).

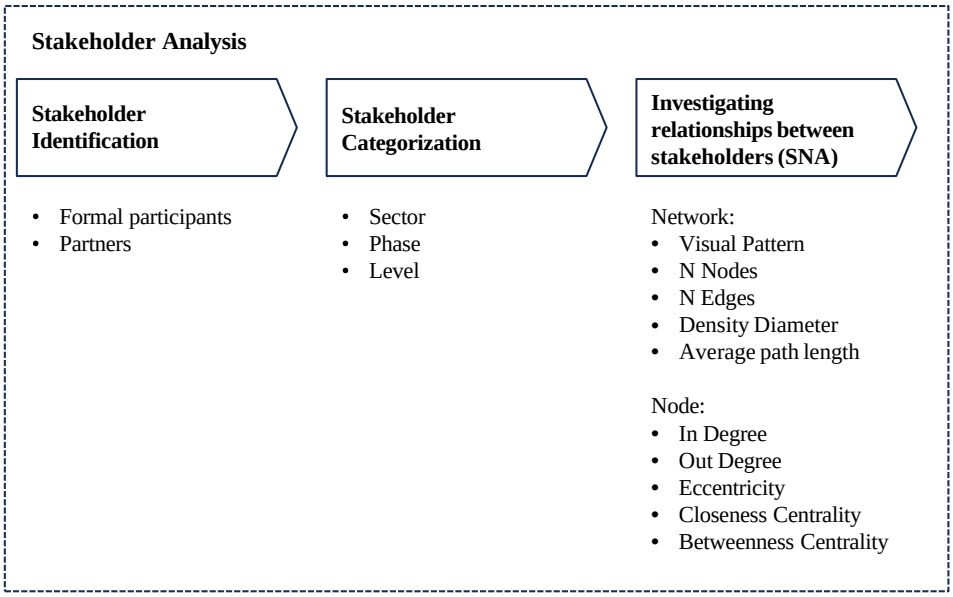


Figure 3. Stakeholder Analysis Method

Adapted from Reed et al., 2009.

Area of Study

The study was conducted as part of a Horizon 2020 project aimed at developing technological solutions to promote holistic wildfire management in the European Union. Technology development was performed in pilot regions where a stakeholder group or organization (henceforth "pilot leader") established a network of collaborators to achieve specific project objectives such as conductingexperiments, exercises, prototype testing, or developing risk management plans. Governmental institutions formally participated in all the pilots from their launch and new actors were engaged during implementation, with the networks growing as the project progressed.

The pilots are being implemented over a 42-month period between December 2022 and May 2026. For the purposes of this study, we selected the pilots with the five largest networks out of the eight project pilots. The pilots selected were in Austria, Greece, Italy, Norway, and Spain. These pilots comprise different scopes, landscapes, weather conditions, socio-economic characteristics, and wildfire management systems.

Table 1. Profiles of the Pilot Cases

Country	Leader Profile	Focus	Scope
Norway	Research Institution	Preparedness for wildfires in northern Europe	Field exercises on forest fires to enhance prevention and preparedness. Test fire protection technologies. Development of guidelines for safety zones and technical requirements.
Italy	Consultancy	Prevention, detection, and adaptation in densely populated areas	Feasibility study for the development of fire-resilient solutions for infrastructure. Analysis of the interactions between wildfire and critical infrastructure. Establish a resilient connection during wildfire occurrences, extending beyond traditional car accessibility.
Spain	Consultancy	Wide area wildfires in high-risk zones	Development of innovative tools for data collection, analysis as, and exploitation. Offer immediate post-fire information, assessing fire severity and recovery. Analysis of the environmental impact of fires in the region.
Austria	Rescue Service Provider	Prevention, protection, and fast response of (peri)-urban wildfire	Fire prevention and a fast response in peri-urban areas. Improve aerial-data use for prediction of fire evolution.
Greece	Research Institution	Human and animal evacuation	Enhance cross-sectorial collaboration and distribution of critical information. Facilitation of more effective rescue operations for humans and animals.

Stakeholder Identification

The first step of the stakeholder analysis process consisted of identifying the stakeholders participating in the local governance wildfire management networks at mapping workshops conducted between September and December 2022. Pilot leaders invited an initial group of stakeholders to workshops aimed at analyzing their local wildfire management context to identify the challenges, needs, and potential contribution of the stakeholders. The list of stakeholders in each pilot was completed through an iterative process with the pilot leaders and participants during the project. Participants finalized the lists after the workshops. The networks documented are unbounded, where membership is not set by a specific affiliation to a policy process, nor to formal engagement in the pilot.

Stakeholder Categorization

Stakeholder categorization aimed to identify the types of stakeholders participating in wildfire management networks. Categorization was performed directly after the initial workshops through a top-down approach where

the study team collaborated with the pilot leaders to identify the stakeholder groups according to three categories:

- Sector (government / civil society / private): categorization based on the governance mode and strategic objectives of each stakeholder group.
- Level (local / regional / national/ international): categorization based on the main level of operation for the stakeholder groups.
- Phase (review / risk reduction / readiness / response / recovery): categorization of phases based on Moore's framework of wildfire management phases (2019) where each stakeholder group is attributed a primary management phase according to their mission.

Stakeholder Relationship Mapping

The third and final step of the stakeholder analysis process involved investigating the relationships among stakeholders through social network analysis. Analysis was divided into two parts: a survey and a workshop. The survey aimed to identify active stakeholders in the region, their relationships, and connections requiring improvement. The workshop was designed to provide a preliminary analysis and validate the survey findings. In this part, stakeholders under the guidance of a pilot leader were asked to visually represent the survey results in a preliminary social network graph.

Data for the surveys and workshops was treated and analyzed in Gephi to create individual network graphs for each area of interest. The analysis process used information from the previous phases to characterize each network according to the attributes in Table 2.

Table 2. SNA variables

Network-level statistics	Node-level statistics
N Stakeholder Groups	In degree
N Connections	Eccentricity
N Stakeholders Groups per sector	Closeness Centrality
N Stakeholders Groups per level	Betweenness Centrality
N Stakeholders Groups per phase	
Network diameter	
Average degree	
Average) Weighted degree	
Average path length	

Ultimately, the study aims to provide empirical evidence to evaluate five European wildfire management contexts through networked governance models by analyzing the configuration of their stakeholder networks. This is measured by assessing network configurations and stakeholders. At the network level, the measurements assess the size, density, and diversity of sectors, in addition to scalar levels of operation and phases. At the stakeholder level, relevance is assessed by measurements such as in degree, eccentricity, closeness centrality , and betweenness centrality. Analysis is completed by a qualitative assessment of the network profile, leading to the discussion on the main factors determining the network configurations and the potential outcomes of these attributes.

RESULTS

Network Configuration Analysis

The first relevant result of the stakeholder analysis process are the profiles of each pilot network in terms of number of stakeholder groups, connections, average degree, network diameter, average path length, density , and the diversity of sectors, levels, and phases of operation. Analysis of network-level statistics allowed an in depth understanding of each network profile and the comparison of results across groups in information provided about the size, the interconnectedness, and the diversity of each pilot network.

The first relevant result of the SNA mapping is that the project is a building a diverse set of networks. The pilot

networks range from complex networks with dozens of stakeholders spanning three or four countries to very localized connections, where a handful of stakeholders interact to conduct specific experiments or test some of the technologies under development. Other attributes vary considerably such as interaction density, the diversity of sectors and scalar levels of operation represented, and the average degree.

Regarding network size statistics, the average number of stakeholders is 46, with the smallest network having 34 and the largest 79. The number of connections ranged from 80 to 295, with an average of 176. The diameter of networks ranges from 3 to 4, with an average of 3.4. The degree of interconnectedness of stakeholders within the networks ranges from 1.66 to 6.97, with an average score across the five pilots of 3.83. The average path length ranges from 1.42 to 2.38, with an average of 1.87. Finally, the graph density ranges from 0.04 to 0.21, with an average of 0.09.

Table 3. SNA Results Overview

Variable	Austria	Greece	Italy	Norway	Spain	TOTAL	AVG
Stakeholders Total	38	38	44	76	34	230	46,00
Connections Total	119	139	80	291	246	875	175,00
Government	28	27	26	40	24	145	29,00
Civil Society	9	4	10	26	7	56	11,20
Private Sector	1	7	8	10	3	29	5,80
Level - Local	20	14	20	15	17	86	17,20
Level - Regional	0	2	8	1	1	12	2,40
Level - National	17	17	15	38	14	101	20,20
Level - International	1	5	1	22	2	31	6,20
Phase - Review	4	10	0	6	2	22	4,40
Phase - Prevention	13	11	23	32	17	96	19,20
Phase - Preparedness	3	5	2	15	9	34	6,80
Phase - Response	18	12	19	22	4	75	15,00
Phase - Recovery	0	0	0	1	2	3	0,60
Average Degree	3,13	3,66	1,66	3,83	6,97	NA	3,85
Avg. Weight. Degree	3,13	4,63	2,59	3,86	10,00	NA	4,84
Network Diameter	3,00	3,00	4,00	4,00	3,00	NA	3,40
Avg. Path Length	1,42	1,87	1,86	2,38	1,81	NA	1,87
Graph Density	0,09	0,10	0,04	0,05	0,21	NA	0,10
Avg Clustering Coef	0,68	0,35	0,31	0,25	0,55	NA	0,43

Data was processed according to the categories to access the diversity of sectors. On average, the pilot networks have 63% of representatives from government, 24% from civil society, and 13% of the private sector. Concentration of actors from government ranges from 74% to 59%. Civil society and private sector participation ranges from 34% to 11% and 18% to 3% respectively. All pilot networks in the study had at least one representative from each sector.

Information on the phases followed a similar pattern, with stakeholders more concentrated, on average, on the phases of response 32% and prevention 42%. Other phases average lower concentrations: review 10%, readiness 15% and recovery 1%. As for the distribution, the pilots that presented the highest concentration in one single phase were the Austrian and Italian pilot networks, with 47% and 43% of its stakeholders concentrated in the response phase respectively. Three out of the five pilot networks did not have stakeholders with priority focus on the recovery phase. Moreover, the Italian pilot did not have stakeholders in the review phase.

Network Composition Analysis.

The investigation of the network attributes of individual nodes complements the analysis across pilots by providing insights on the most well-connected types of stakeholders in each pilot, along with the identification of gaps identified by poorly connected stakeholders in key areas.

The first analysis identified of the most relevant stakeholder in each pilot according to the number of connections received. An investigation of the stakeholders with the highest in-degree scores indicates that NGV43 (Norway, 21), SGV27 (Spain, 17), DD Chania (Greece, 16), and IGV2 (Italy, 6) are the actors receiving the highest number of connections. In Austria, the AGV8, AGV13, ACS16 and AGV19 are tied for the most connections with five each. All stakeholders with the highest in-degree scores are from government, except for ACS16, which is a civil society entity providing public services. In terms of the phase focus, four out of five of the most highly ranked stakeholders perform functions in the "Risk Reduction – Prevention" phase, while one is in "Risk Reduction – Readiness" and two in "Response - Fire Fighting Operations."

The second analysis identified the stakeholder most strategically connected with betweenness centrality scores. Betweenness centrality is a measure of the dependency of other nodes on a given node (Brandes et al., 2016) and refers to the number of times a given node appears in the optimal connection between two other nodes in the graph. In our analysis, the pilot leaders in Spain (SGV14, 191), and Italy, (IGV2, 152) had the highest scores. In Austria, ACS10 (36), a stakeholder that specializes in networking for the improvement of wildfire management had the highest score. In Greece, HGV13 (116), a governmental agency responsible for coordinating emergency response, and in Norway, NCS47 (1013), a research center that partners with the pilot leaders, had the highest scores. In all cases the highest scores are attributed to actors that are central to the pilot's operation or to the management of networks of wildfire management in the implementation region.

A final investigation into the stakeholders' strategic placement was performed through closeness centrality measures. The results from this variable also indicate the strategic placement of the stakeholders in relation to the network but using its distance to all other nodes as a measure of interconnectivity. Here the lowest values are attributed to the most well-connected stakeholders, with only stakeholders that actively responded to the survey considered in the analysis. Of note, the highest scores in closeness centrality are attributed to stakeholders that might not be central to the network in terms of in-degree and out-degree scores, but are directed connected to them, as the case of SCS2 in Spain (0.40), HGV15 in Greece (0.42), ICS1 in Italy (0.46), AGV7 in Austria (0.49), and NGV64 (0.33) in Norway. These best scoring cases showcase the benefits of reaching out to network brokers not only to the performance of everyday duties but also to the development of integrated wildfire management actions.

Norway

The Norwegian pilot is led by a Norwegian center with expertise in fire resistance, reaction to fire, firefighting, fire tests, and fire simulations. The pilot conducted a hybrid SNA workshops, with stakeholders completing the survey prior to the on-site event in September 2023. The SNA survey yielded 30 responses resulting in the identification of 79 different stakeholders. The network of the Norwegian pilot is diverse and extensive , involving 44 stakeholders from the government, 25 from civil society, and 10 from the private sector. The stakeholders represent a broad spectrum, ranging from government bodies (e.g., local and regional authorities, fire brigade departments, and directorates) and educational institutions to agricultural associations and fire research organizations. The private sector also has a notable presence among the stakeholders, representing forest owner associations, technology development, and manufacturing companies, among others.

Connections within the network are intricate, linking entities involved in wood processing, fire safety, agriculture, meteorology, research, and technology innovation. Notably, the network includes international associations and research institutions, European agencies, and global non-profits, showcasing a comprehensive and multi-dimensional approach. The network appears to be well balanced, with various stakeholders able to contribute their expertise and resources to address fire-related issues comprehensively. This inclusive approach involves key players from different sectors, emphasizing a holistic strategy in managing and preventing wildfires.

The Norwegian pilot's activities were conducted in forested inland areas in the east and coastal landscapes with heather, grass, and scrub vegetation in the west. The exercises and experiments are designed and implemented in coordination with local authorities.

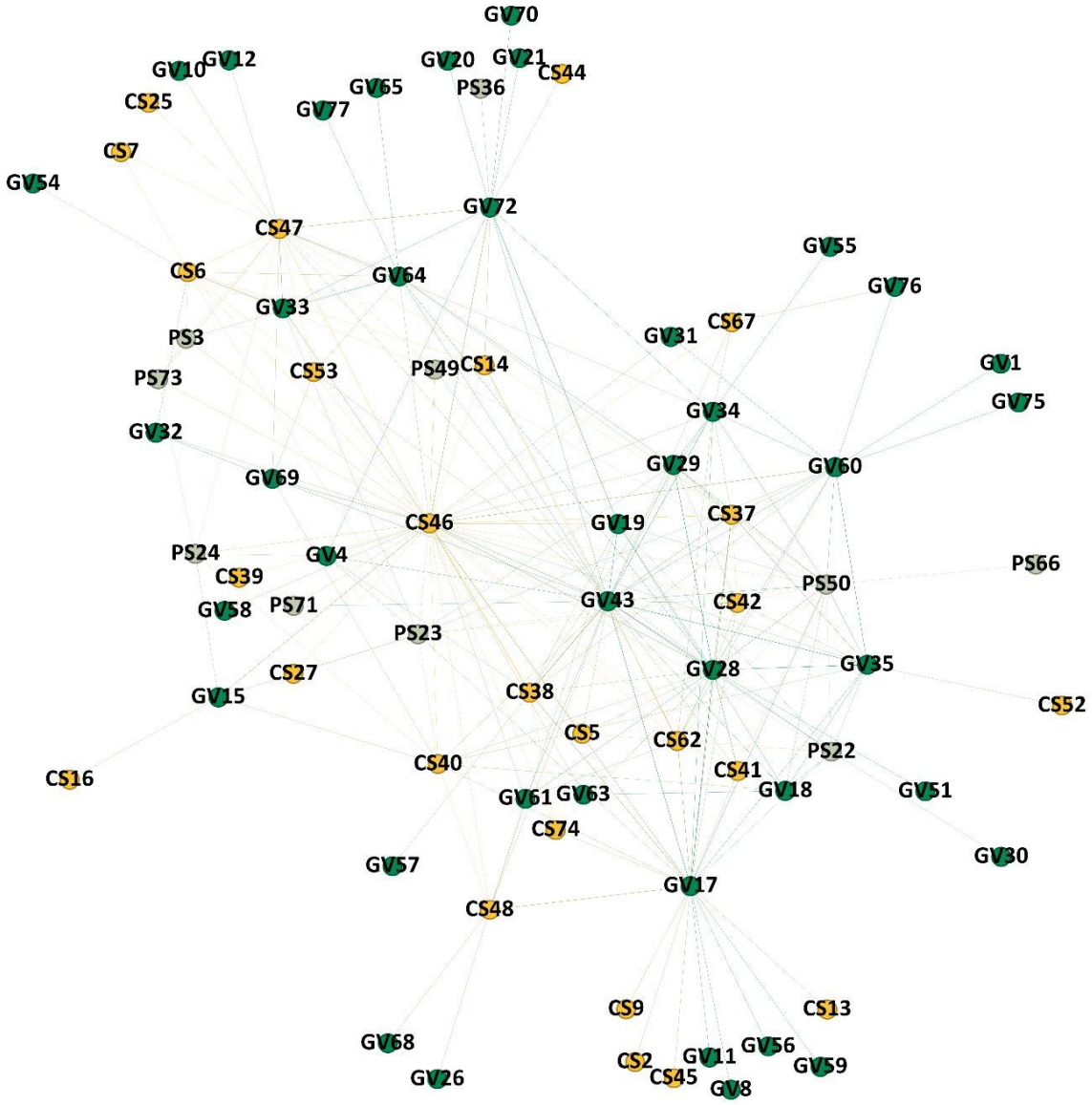


Figure 4. Norway Stakeholder Network

Italy

The Italian pilot is led by a private non-profit company engaged in research activities and the experimental demonstration of innovative technologies for the construction sector. The pilot leader organized an online SNA workshop in May 2023. The SNA survey was completed by 6 respondents from different organizations, resulting in the identification of 44 stakeholders.

The Italian network of stakeholders exhibits a comprehensive structure with robust connections to regional and local authorities. This is evident by the significant involvement of stakeholders from municipal, regional, and community levels. The identification of 26 stakeholders from government and 10 from civil protection emphasizes the active involvement of local entities in addressing wildfire management. The modest representation of private sector stakeholders (8) reinforces the idea that the predominant role in wildfire management is played by governmental and civil protection bodies. Such an approach emphasizes the vital role of community participation and collaboration at the local level in effectively managing and mitigating the impact of wildfires.

The Italian pilot focuses on the Sorento Peninsula, a densely populated area with urban areas and forests on the surrounding slopes.

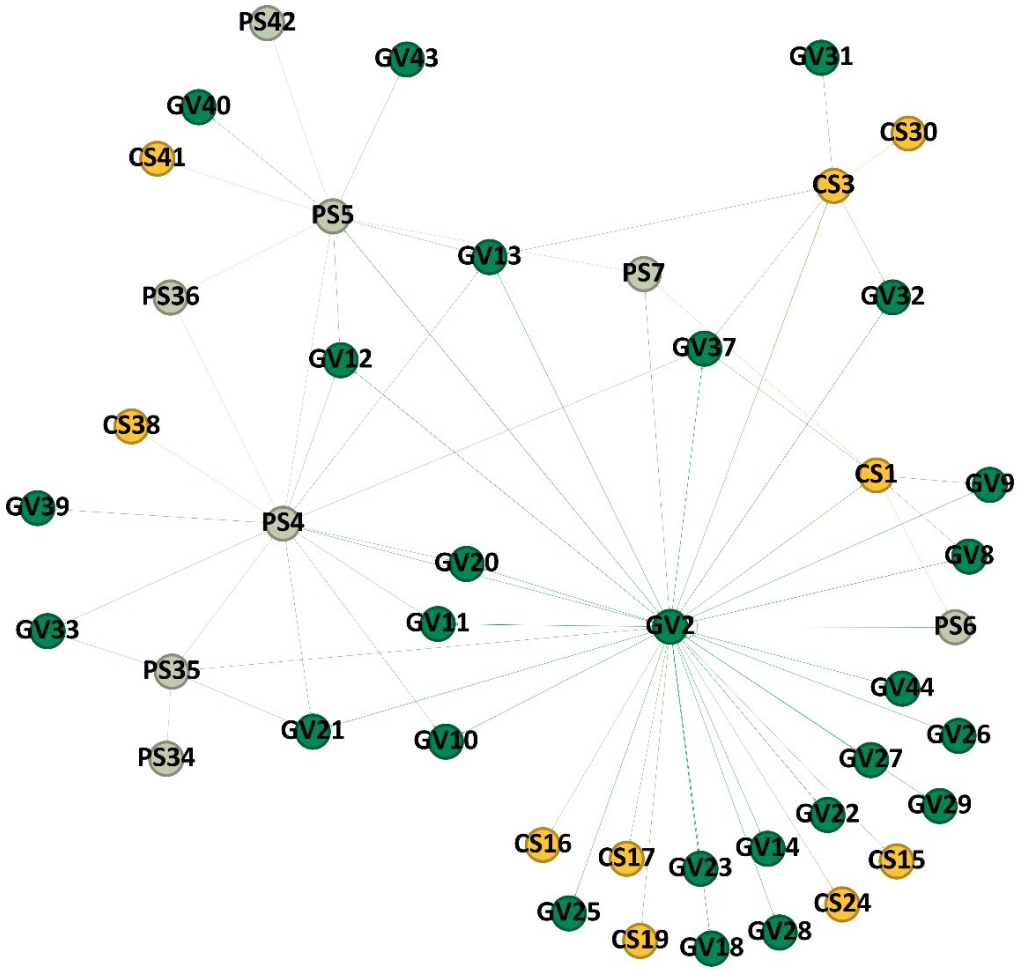


Figure 5. Italy stakeholder network

Spain

The Spanish pilot is led by an innovation and high-tech engineering consulting company. The SNA survey conducted in October 2023 had 21 respondents that identified 34 stakeholders across a diverse range.

The stakeholder types included 24 types of government entities, 3 from the private sector, 7 from the civil society. The governmental stakeholders appear to be the most common in the pilot network, including different municipalities of the autonomous community of Castile and León, and state environmental and forest protection services. Civil society is represented by environmental agencies and associations, forest management advisors, and forest owner associations. Lastly, the private sector brings stakeholders to the technological, innovation, and research fields.

While there is a substantial government presence, particularly of local authorities, firefighting, and forest management services, the involvement of civil society organizations, ecological associations, and private companies like Capgemini indicate a distributed network. This mix suggests a more balanced strategy rather than a purely top-down approach. The pilot is set to investigate the wildfire-prone Pinus species and mixed forest, recognized as high risk.

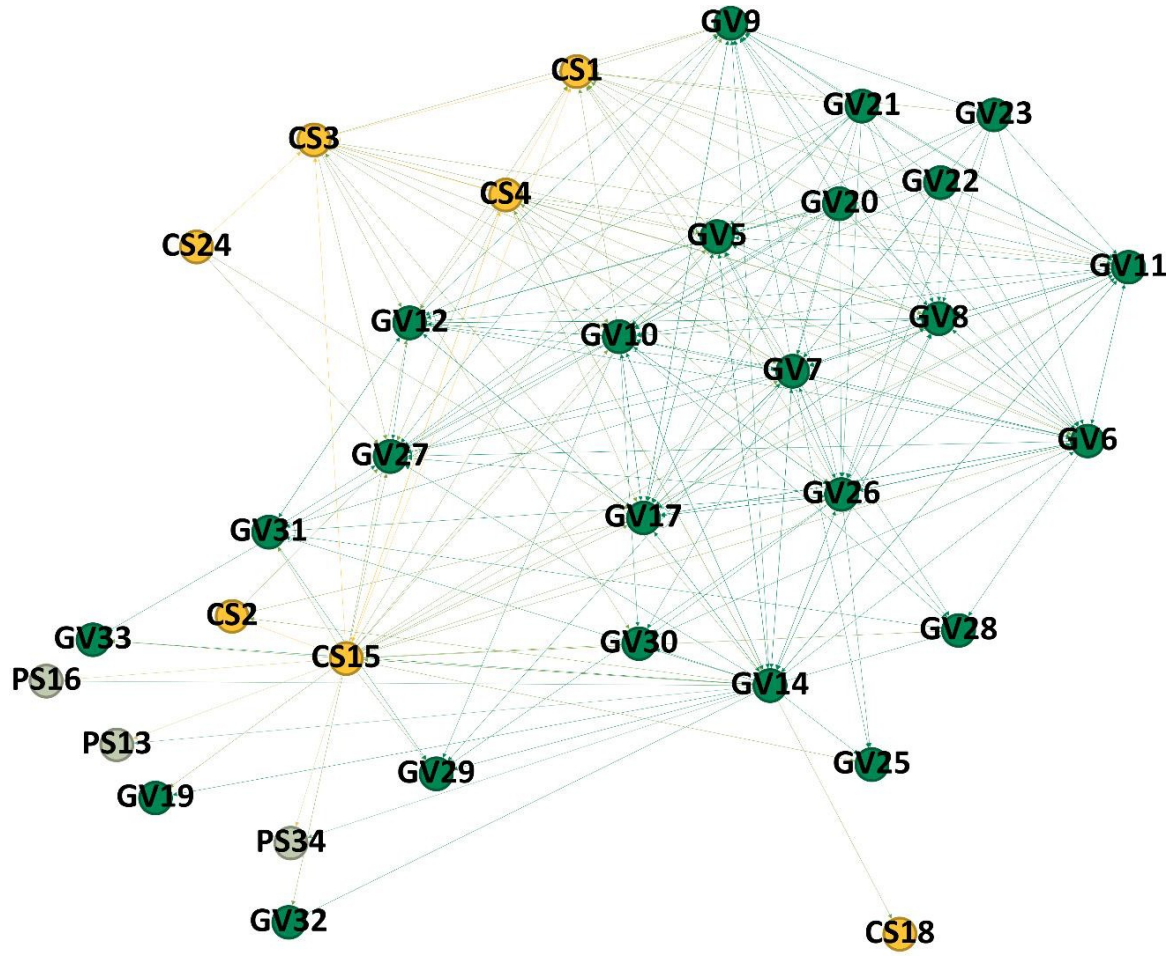


Figure 6. Spanish Stakeholder Network

Austria

The Austrian pilot is led by an organization that emphasizes health-related education and training. The pilot team

held a physical SNA workshop in October 2023. The SNA survey, conducted with 5 respondents, identified a total of 38 stakeholders, revealing a network predominantly composed of 28 government entities, 9 from civil society, and only 1 from the private sector.

Analysis reflects a blended approach, characterized by a significant top -down influence from governmental entities, coupled with engagement from civil society, but limited participation from the private sector. The pilot is set to evaluate the tools and solutions provided by the technical partners under realistic conditions and data with the end-users.

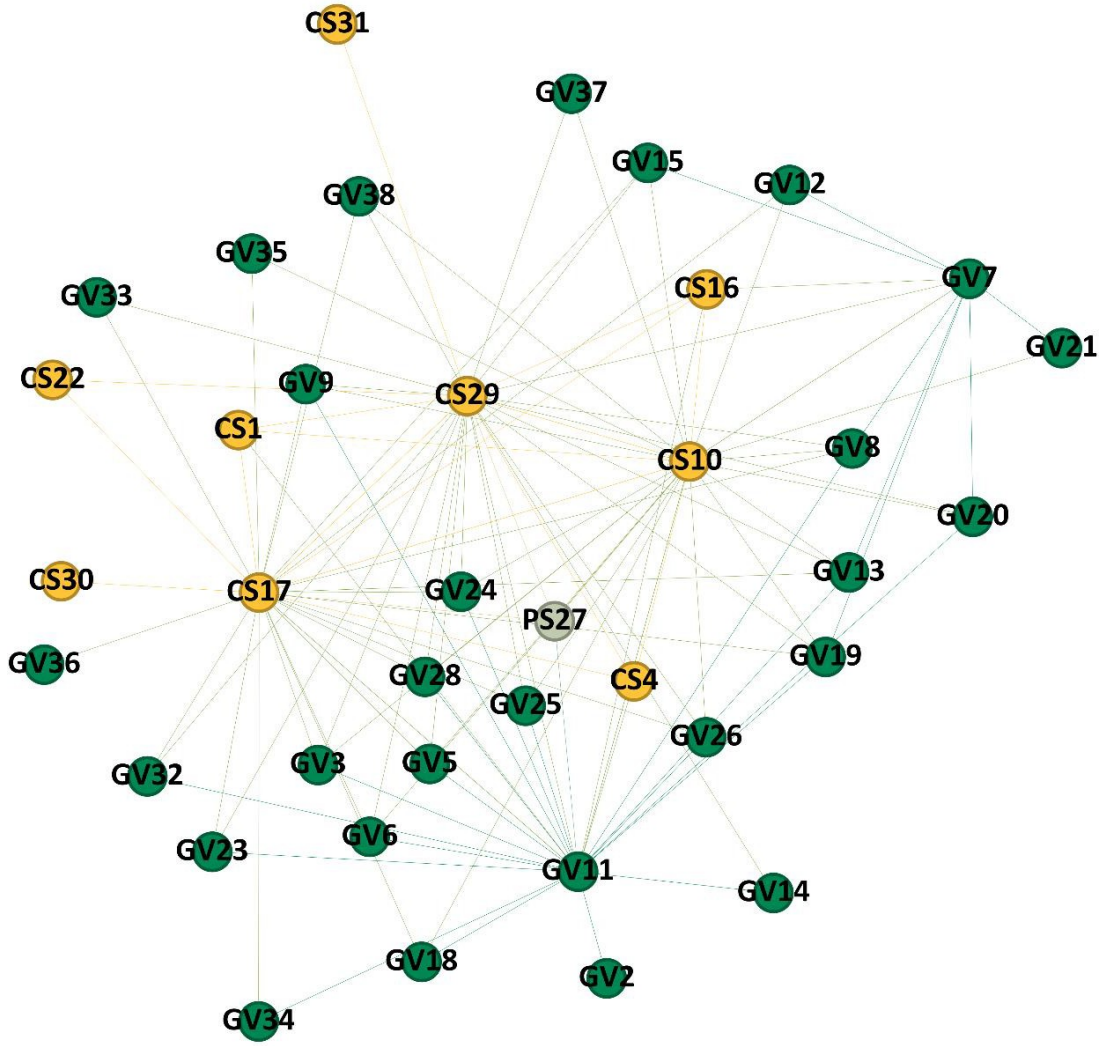


Figure 7. Austria Stakeholder Network

Greece

The Greek pilot is led by a university investigating wildfire management in the challenging terrain of the Samaria Gorge. The university organized a physical SNA workshop in October 2023. 15 participants completed the SNA survey, resulting in the identification of 38 stakeholders. The government sector, counting 27 stakeholders, takes a central position, involving fire brigade departments, local authorities, and public research centers, alongside entities like Police and forest services. The private sector is represented by 7 organizations, including research and innovation centers, and tourist and hospitality local companies that contribute cutting-edge technologies. 4 types of stakeholders represent civil society, with humanitarian organizations and mountaineering clubs in the network. The analysis reveals a harmonious top-down and bottom-up approach.

The use case is in southwest Crete in the regional unit of Chania, a closed canyon that limits the applicability of standard procedures.

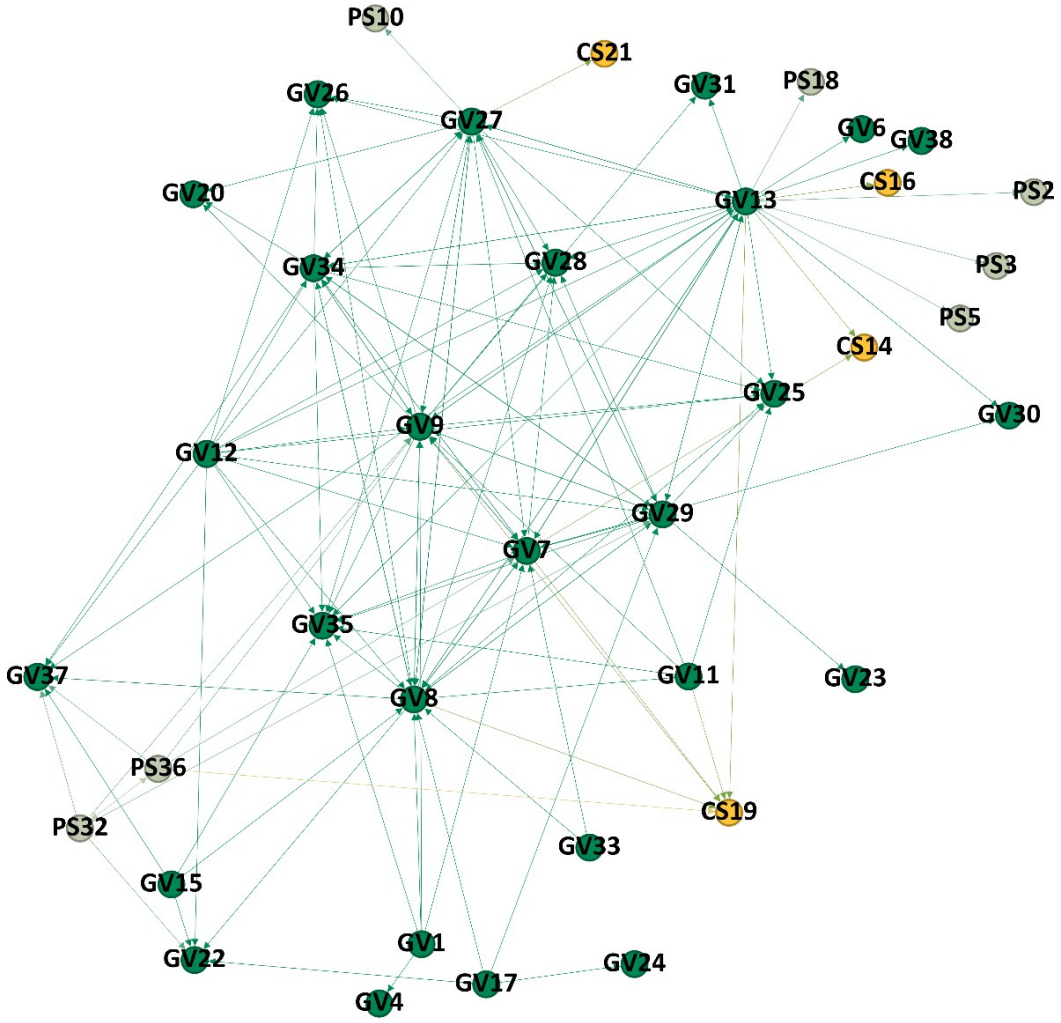


Figure 8. Greece Stakeholder Network

CONCLUSION

The project is designed to implement a holistic wildfire management approach. In practical terms, the approach involves the development of solutions across the wildfire management cycle and the coordination among different stakeholders and sectors. By profiling each of the eight pilots, the SNA first illustrates how this coordination is carried out. The scope of action, the profile of stakeholders and the density of connections are seen as evidence of this engagement.

The SNA analysis also identifies the key stakeholders in each pilot and provides insights on the connection gaps. The identification of these gaps provides two immediate inputs for the project: it supports local engagement in the pilot regions, and it assesses fit between technologies and local challenges.

In the wildfire management and disaster response policy debate, the analysis adds to the few empirical studies on natural resource governance networks in the European Union. It is relevant that the study is being performed while the project is under implementation, which allows the results to be shared with other similar EU fire projects and officials in the European commission. The results contribute to the fields of wildfire studies and networked governance by assessing how efforts such as the project contribute to improving governance in the territories of implementation.

REFERENCES

AGIF. (2023*). Landscape Fire Governance Framework 2023* (pp. 1–20). Portuguese Agency for Integrated Rural Fire Management. https://www.wildfire2023.pt/conference/framework

Battistoni, P., Di Gregorio, M., Giordano, D., Sebillo, M., Tortora, G., and Vitiello, G. (2020). Wearable Interfaces and Advanced Sensors to Enhance Firefighters Safety in Forest Fires. *In Proceedings of the International Conference on Advanced Visual Interfaces*, Salerno Italy: ACM, 1–3. doi:10.1145/3399715.3399961.

Carrington, P. J., Scott, J., & Wasserman, S. (Eds.). (2005). *Models and methods in social network analysis.* Cambridge University Press.

Chuvieco, E., Yebra, M., Martino, S., Thonicke, K., Gómez-Giménez, M., San-Miguel, J., Oom, D., Velea, R., Mouillot, F., Molina, J. R., Miranda, A. I., Lopes, D., Salis, M., Bugaric, M., Sofiev, M., Kadantsev, E., Gitas, I. Z., Stavrakoudis, D., Eftychidis, G., … Viegas, D. (2023). Towards an Integrated Approach to Wildfire Risk Assessment: When, Where, What and How Ma y the Landscapes Burn. *Fire*, 6(5), Article 5. https://doi.org/10.3390/fire6050215

Edgeley, C. M., & Paveglio, T. B. (2016). Influences on stakeholder support for a wildfire early warning system in a UK protected area. *Environmental Hazards*, 15(4), 327–342. https://doi.org/10.1080/17477891.2016.1209155

Ferrara, C., Salvati, L., Corona, P., Romano, R., & Marchi, M. (2019). The background context matters: Local-scale socioeconomic conditions and the spatial distribution of wildfires in Italy. *Science of the Total Environment*, 654, 43–52. https://doi.org/10.1016/j.scitotenv.2018.11.049

Giuliani, E., & Bell, M. (2005). The micro-determinants of meso-level learning and innovation: Evidence from a Chilean wine cluster. *Research Policy*, 34(1), 47–68. https://doi.org/10.1016/j.respol.2004.10.008

Grimble, R., & Wellard, K. (1997). Stakeholder methodologies in natural resource management: A review of principles, contexts, experiences and opportunities. *Agricultural Systems*, 55(2), 173–193. https://doi.org/10.1016/S0308-521X(97)00006-1

Hamilton, M., Nielsen-Pincus, M., & Evers, C. (2023). Wildfire risk governance from the bottom up: Linking local planning processes in fragmented landscapes. *Ecology and Society*, 28(3), art3. https://doi.org/10.5751/ES-13856-280303

Hauck, J., Schmidt, J., & Werner, A. (2016a). Using social network analysis to identify key stakeholders in agricultural biodiversity governance and related land-use decisions at regional and local level. *Ecology and Society*, 21(2). https://doi.org/10.5751/ES-08596-210249

Hauck, J., Schmidt, J., & Werner, A. (2016b). Using social network analysis to identify key stakeholders in agricultural biodiversity governance and related land-use decisions at regional and local level. *Ecology and Society*, 21(2), art49. https://doi.org/10.5751/ES-08596-210249

Kirschner, J., Clark, J., & Boustras, G. (2023). Governing wildfires: Toward a systematic analytical framework. *Ecology and Society*, 28(2), art6. https://doi.org/10.5751/ES-13920-280206

Krupa, M., Cenek, M., Powell, J., & Trammell, E. J. (2018). Mapping the Stakeholders: Using Social Network Analysis to Increase the Legitimacy and Transparency of Participatory Scenario Planning. *Society & Natural Resources*, 31(1), 136–141. https://doi.org/10.1080/08941920.2017.1376140

Lecina-Diaz, J., Campos, J., Pais, S., Carvalho-Santos, C., Azevedo, J., Fernandes, P., Gonçalves, J., Aquilué, N., Roces-Díaz, J. V., Agrelo De La Torre, M., Brotons, L., Chas-Amil, M.-L., Lomba, A., Duane, A., Moreira, F., Touza, J., Hermoso, V., Sil, Â., Vicente, J., … Regos, A. (2023). Stakeholder perceptions of wildfire management strategies as nature-based solutions in two Iberian biosphere reserves. *Ecology and Society*, 28(1), art39. https://doi.org/10.5751/ES-13907-280139

Marchi, M., Chianucci, F., Ferrara, C., Pontuale, G., Pontuale, E., Mavrakis, A., Morrow, N., Rossi, F., & Salvati, L. (2018). Sustainable land-use, wildfires, and evolving local contexts in a Mediterranean Country, 2000-2015. *Sustainability*, 10(11). https://doi.org/10.3390/su10113911

Moore, P. F. (2019). Global Wildland Fire Management Research Needs. *Current Forestry Reports*, 5(4), 210–225. https://doi.org/10.1007/s40725-019-00099-y

Provan, K. G., & Kenis, P. (2007). Modes of Network Governance: Structure, Management, and Effectiveness. *Journal of Public Administration Research and Theory*, 18(2), 229–252. https://doi.org/10.1093/jopart/mum015

Reed, M. S., Graves, A., Dandy, N., Posthumus, H., Hubacek, K., Morris, J., Prell, C., Quinn, C. H., & Stringer, L. C. (2009). Who's in and why? A typology of stakeholder analysis methods for natural resource management. *Journal of Environmental Management*, 90(5), 1933–1949. https://doi.org/10.1016/j.jenvman.2009.01.001

Roberts, R. M., Jones, K. W., Duke, E., Shinbrot, X., Harper, E. E., Fons, E., Cheng, A. S., & Wolk, B. H. (2019). Stakeholder perceptions and scientific evidence linking wildfire mitigation treatments to societal outcomes. *Journal of Environmental Management*, 248, 109286. https://doi.org/10.1016/j.jenvman.2019.109286

Steelman, T. (2016). U.S. wildfire governance as social-ecological problem. *Ecology and Society*, 21(4), art3. https://doi.org/10.5751/ES-08681-210403

Steelman, T., Nowell, B., Velez, A.-L., & Scott, R. (2021). Pathways of Representation in Network Governance: Evidence from Multi-Jurisdictional Disasters. *Journal of Public Administration Research and Theory*, 31(4), 723–739. https://doi.org/10.1093/jopart/muab004

Thompson, M. P., MacGregor, D. G., Dunn, C. J., Calkin, D. E., & Phipps, J. (2018). Rethinking the Wildland Fire Management System. *Journal of Forestry*, 116(4), 382–390. https://doi.org/10.1093/jofore/fvy020

UNISDR. (2017). *Technical Guidance for Monitoring and Reporting on Progress in Achieving the Global Targets of the Sendai Framework for Disaster Risk Reduction*. United Nations. https://www.undrr.org/publication/technical-guidance-monitoring-and-reporting-progress-achieving-global-targets-sendai

United Nations. (2015). *Sendai Framework for Disaster Risk Reduction 2015—2030*. United Nations. https://www.preventionweb.net/files/43291_sendaiframeworkfordrren.pdf?_gl=1*goneqo*_ga*MzczNTAzNDcwLjE2OTM5OTE1NzM.*_ga_D8G5WXP6YM*MTY5NTI4MDU5OC4yLjAuMTY5NTI4MDU5OC4wLjAuMA..

Valeri, M., & Baggio, R. (2021). Italian tourism intermediaries: A social network analysis exploration. *Current Issues in Tourism*, 24(9), 1270–1283. https://doi.org/10.1080/13683500.2020.1777950

Wagner, K. R. H. (2022). Designing insurance for climate change. *Nature Climate Change*, 12(12), 1070–1072. https://doi.org/10.1038/s41558-022-01514-2

Wasserman, S., & Faust, K. (1994). *Social Network Analysis_ Methods and Applications (Structural Analysis in the Social Sciences)*. Cambridge University Press.



Wildfire Evacuation: Social Resilience and Local Community Response

Track Chairs

Max Kinateder, NRC Canada

Paul Gegerg, Vereinigung zur Förderung des Deutschen Brandschutzes e.V.

Steve Gwynne, Movement Strategies

Noureddine Bénichou, NRC Canada

Evacuation modelling for community resilience: An Australian perspective

Dhirendra Singh*
CSIRO's Data61

Hanna Grzybowska
CSIRO's Data61

Leorey Marquez
CSIRO's Data61

Vincent Lemiale
CSIRO's Data61

ABSTRACT

Communities worldwide are increasingly facing severe impacts from natural hazards like wildfires, floods, and cyclones. Where a life-threatening hazard is imminent, and moving people is a safe option, large-scale evacuations of communities are often undertaken by authorities. In Australia, where evacuation is a recommendation to individuals and households, and not an enforceable order, community safety requires a collaborative approach: where communities take ownership of the risk and responsibility for their actions, and work closely with emergency management agencies for planning, preparedness, readiness, and response efforts. In this context, drawing examples from existing works, we present requirements for evacuation modelling from the perspectives of different end-users (the at-risk individuals and households, the town planner and developer, the risk analyst, the incident controller and response crew). We argue that this technology is useful, beyond its obvious use in planning and decision support, in building shared understanding among all stakeholders for improving community safety.

Keywords

Wildfires, Evacuation, Modelling, Simulation, Resilience

INTRODUCTION

With the increasing frequency and severity of extreme weather-induced natural hazards worldwide, large-scale evacuations of communities out of harm's way are becoming increasingly regular. Between 2017-19, in excess of one million people were evacuated from wildfires in California, USA (Wong & Broader, 2020). In 2023 alone, an estimated 200,000 people were evacuated during Canada's unprecedented fire season (Beverley, 2023), more than the combined total for the 1979-2007 period (Beverly & Bothwell, 2011). Hundreds of *bushfires*, the analogous term for wildfires in Australia, devastate vast landscapes in Australia every year, and can bring with them disastrous consequences for communities (Binskin et al., 2020; Teague et al., 2010). Against this backdrop, understanding evacuation-related human behaviour to inform choices in future events is becoming vital for all involved, from evacuating communities to emergency management agencies tasked with their timely egress to safety.

Recent years have seen significant advances in computer-based simulation technology for modelling large-scale evacuations, offering many benefits to planning and preparedness for future events. In the simplest use case, this technology can help improve understanding of traffic and timing when moving populations en masse (Goto et al., 2012; Wahlqvist et al., 2021). Combined with hazard progression modelling, it can allow evacuation policies to be stress tested under "what-if" scenarios (Goto et al., 2012; Wahlqvist et al., 2021), inform training of emergency personnel (Ronchi et al., 2017), improve community education and information (Goto et al., 2012), and bring a shared understanding of roles and responsibilities between authorities and communities (Chen et al., 2021). The ability to model complex human behaviour at a granular level, even down to every individual in a community, using techniques like agent-based modelling (Goto et al., 2012; Wahlqvist et al., 2021), are nothing short of impressive (Goto et al., 2012; Wahlqvist et al., 2021), and work to demonstrate that the technology is up to the challenge and abounds in academia (Ronchi et al., 2017).

*Corresponding author. Authors' emails follow the format firstname.lastname@data61.csiro.au.

Table 1. Evacuation modelling use cases.

	Planning	Preparedness	Readiness	Response
Decision timeline	Years ahead	Season ahead	Week ahead	Day of incident
Use cases	Integrated land use and transport planning, Risk reduction planning, Emergency planning for townships, Community education	Seasonal risk assessment and reduction, Adapting emergency plans, Preparedness exercises, Community information, Special events (e.g., festivals)	Emergency resources allocation, Community messaging, Special events arrangements	Evacuation options and messaging
Example actions	Constructing and upgrading community refuges, Identifying and upgrading roads and bridges for safe egress, Establishing standard operating procedures, roles and responsibilities	Pre-season incident management training, Community awareness programs, Adapting township emergency plans, informing traffic management arrangements	Community information, Leave early messaging, Event management, traffic management, emergency resources positioning and readiness	Evaluating if/when/how evacuations are conducted, Timing of leave or shelter messaging, Providing directions for safe egress, Placement of traffic management resources
End users	Agencies responsible for risk reduction, town and infrastructure planning, emergency management, Local / State / Federal governments, Building & construction industry	Emergency management agencies, Local / State governments, Township emergency planning committees, Community groups, Local residents, Event planners, Tourism agencies	Emergency management agencies, Community groups, Local residents, Event planners, Tourism agencies	Emergency management agencies

Despite improvements in computing software and hardware that have made evacuation simulation technology accessible today, these models continue to live largely in the domain of researchers and not practitioners. This sits in stark contrast to models of natural hazards, such as bushfires (Bally, 2023; Faggian et al., 2017), that are widely used in emergency management at all levels of operations and across all timescales from planning to response. Why so? In this paper we argue that the uptake of evacuation models is hindered today not as much due to technological challenges as it is due to social ones. What is lacking often is a clear understanding of the audience and a solution adapted for its needs. *Evacuation modelling for whom?* We argue for the need of a socio-technical approach: one that optimises not just the technological solution, but equally, and jointly, its use in society. In this setting there is no linear technology transfer from producer to consumer, rather an iterative process of co-production, co-learning, and product evolution.

While it is evident that evacuation related decisions must be made at the time of a bushfire, the timeline for evacuation decisions begins several years prior to the event. For many jurisdictions, emergency management arrangements relate to the timing of decisions before, during and after an emergency, and apply to the four broadly defined phases of Prevention, Preparedness, Response, and Recovery (PPRR) (Community Outcomes and Recovery Subcommittee (CORS), 2022). In this framing, evacuation decisions fall under the first three phases. When assessing how evacuation modelling can help inform decisions, it is useful to consider *when* decisions have to be made, *what* decisions can benefit from the outputs of the modelling, and *who* is making the decisions. Table 1 shows the timeline before and during a bushfire split into four phases: Planning that begins several years before the event and where related actions can take years to implement, Preparedness where actions align with the cycle of seasons every year, Readiness defined by actions in the days leading to an event, and Response relating to actions during the event. In this paper, and with a focus on Australia, we argue the need for a deeper appreciation of these use cases and end users and their requirements, when building evacuation modelling and simulation capability.

Consideration of community evacuations is a relatively new concern in Australia. As recently as the disastrous 2009 Black Saturday bushfires during which 173 lives were lost, Australia had no standing policy around evacuations. The subsequent 2009 Victorian Bushfires Royal Commission (Binskin et al., 2020) made 67 recommendations, of

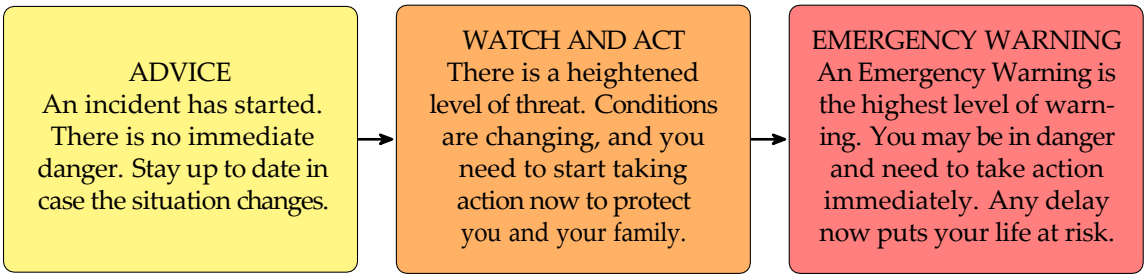


Figure 1. Warning levels under the new Australian Warning System.

which Recommendation 1, 3, and 5, sought a comprehensive approach to evacuations. A decade later, following the devastation of the Black Summer bushfires of 2019-20 that resulted in 34 direct casualties, and an estimated additional 417 deaths from bushfire smoke exposure (Borchers Arriagada et al., 2020), the Royal Commission into National Natural Disaster Arrangements (Binskin et al., 2020) expanded on the need for a national approach to evacuation planning and shelters. The evacuation context in Australia is one of vast landscapes prone to natural hazards and large numbers of relatively small and dispersed populations at their interface. During hazard events, such as bushfires that burn hundreds of thousands of hectares every year in Australia, limited resources exist within emergency services to attend to everyone in need. For the first few critical hours of an emergency until help arrives, it is the collective actions of impacted communities that have the biggest impact on lives saved.

Evacuations in Australia are additionally not mandatory. For this reason, improving community preparedness to take preventative action has been a focus of media campaigns from emergency agencies. Between 2009 and 2020, the nationally agreed Prepare Act Survive campaign saw the percentage of people who were prepared with a plan of action improve from 30 percent to 80 percent (NSW Rural Fire Service, 2023). However, residents who intend to leave have little in the way of knowing what conditions will be like once they become mobile. Evacuation simulation tools with map-based visualisations of expected behaviours and estimated traffic can show how decisions and actions of residents and emergency services translate to traffic and congestion on roads, and help improve individuals' understanding and planning.

The remainder of this paper is structured as follows. In the next section we describe important use cases and end users of evacuation modelling, focusing on Australia, then set out what we believe are key requirements for evacuation modelling capability from an applied perspective towards building understanding and improving community resilience. The subsequent section then covers current state-of-the-art in evacuation modelling capability in this context. Finally, in the discussion we reflect on what is required to get evacuation modelling tools to a level where they can confidently be used across bushfire planning, readiness, and response in Australia.

EVACUATION MODELLING FOR WHOM?

In this section we discuss evacuation modelling from an end user perspective across the decision timeline, in an attempt to highlight the importance of accounting for these perspectives when designing and implementing evacuation modelling technology.

The At-Risk Individual

The new Australian Warning System applies a nationally consistent approach to information and call to action for hazards like bushfire, flood, storm, cyclone, extreme heat and severe weather [1], and was established following recommendations from the Royal Commission into National Natural Disaster Arrangements (Binskin et al., 2020). Under this scheme, three warning levels communicate increasing risk to communities, as shown in Figure 1. Each warning level is associated with a set of action statements that give the community clear advice about what to do. Evacuation-related actions can accompany a Watch and Act warning with action calls including "Prepare to evacuate", or "Evacuate now," while for the more severe Emergency Warning the wording can include "Evacuate immediately" and "Seek shelter now".

An individual's awareness of the situation is a function of many factors, emergency warnings being one. Other important sources that contribute to awareness include visual cues such as the sight of smoke, embers, flames, and neighbours leaving, the sound of emergency vehicles and aircraft, communications from family and friends, and broadcasts through public information channels (Whittaker, 2019). Acknowledgement of the situation at a level that

1https://www.australianwarningsystem.com.au



Figure 2. Examples of the use of evacuation modelling for community information and education as shared by the Commonwealth Scientific and Industrial Research Organisation (CSIRO Evacuation Modelling Group, 2023).

warrants deliberation happens over time and via reinforcement of information from different sources. Individuals vary in their perceptions of risk which influences if and when they take action (McLennan et al., 2019).

Whether situation-aware individuals choose to evacuate varies significantly still. Extensive work by Strahan et al. on self-evacuation archetypes in Australian bushfires (K. Strahan et al., 2018; K. W. Strahan et al., 2019) shows that behaviours of resident individuals can be classified into one of seven distinct groups. A Threat Denier pays little heed to warnings because of the belief that their personal safety is not threatened by the bushfire, making evacuation unnecessary. Similarly, an Experienced Independent is also committed to remaining at home but for the reason that they are experienced and believe they are well prepared to fight the fire. In contrast, a Worried Worried Waverer hesitate between leaving and staying as they are not confident in their ability to defend their property despite prior training. A Community Guided person is constantly seeking information to stay informed, and values official advice when deciding to evacuate. A Responsibility Denier on the other hand is a committed evacuator who expects to be told what action to take when, effectively taking no responsibility for their own safety. Whereas, a Dependent Evacuator is incapable of taking protective action and relies on others for their safe evacuation. Finally, a Considered Evacuator is strongly committed to evacuate as soon as they become aware of the threat. While these self-evacuation archetypes give useful insights into behaviours of residents, they currently do not extend to households or seasonal visitors.

Even where a decision to leave or shelter nearby is made, there is still further variation in how decisions translate to trips. First, the distribution of departure times is skewed with a large number of people typically leaving sooner than the mean departure time, and a relatively smaller proportion leaving quite a long time after. Then a significant proportion of the population may undertake *intermediate* trips, which are trips to destinations other than the final evacuation destination, such as trips to pick up dependents from school or pets from homes. Finally, any trip, including the one to the evacuation destination may be disrupted by a change in the situation such as fire ahead, fallen tree over road, an accident, or redirection from traffic management authorities. In such situations, people may look for alternative routes to their destination, or failing that, look for alternative destinations. Lack of data on evacuation related behaviour and mobility are open challenges for evacuation modelling (Kuligowski, 2021).

Evacuation modelling capability should allow for heterogeneity and variation in behaviour expected during a real event towards more realistic estimates of egress times (Bulumulla et al., 2017; Robertson et al., 2021; Singh et al., 2021) and importantly of risk to communities. Simulations of evacuation can be extremely beneficial for community information, awareness, engagement, preparedness, and ultimately resilience to bushfire threat. Examples of this use case in Australia are seen in the evacuation modelling work being undertaken at the Commonwealth Scientific and Industrial Research Organisation (CSIRO). These include the use of evacuation modelling for information and education at community events (Atkin, 2023; Gamage, 2022, 2023), and for improving resilience through direct community engagement in the co-production of evacuation scenarios with support from the emergency services (Dennis, 2022) (see Figure 2).

The Town Planner and Builder

A recent high-profile case in Western Australia demonstrates the relevance of evacuation modelling in assessing the suitability of proposed new housing developments in high bushfire risk areas. Perth Hills in Western Australia suffered significant bushfires in 2011 (71 homes destroyed) (Keelty, 2011) and 2014 (48 homes destroyed) (Smith et al., 2015). Recently, concerns over a planning proposal for a township to accommodate approximately 1000 new dwellings there has seen significant government scrutiny (Government of Western Australia, 2023) and public protest (Perth Hills Action Group, 2023). The proponent developer submitted updated plans to address concerns, including a report on evacuation modelling for the proposed township. After final review, and hearing from public and experts, the planning committee recommended to the planning commission that the proposal be refused. In

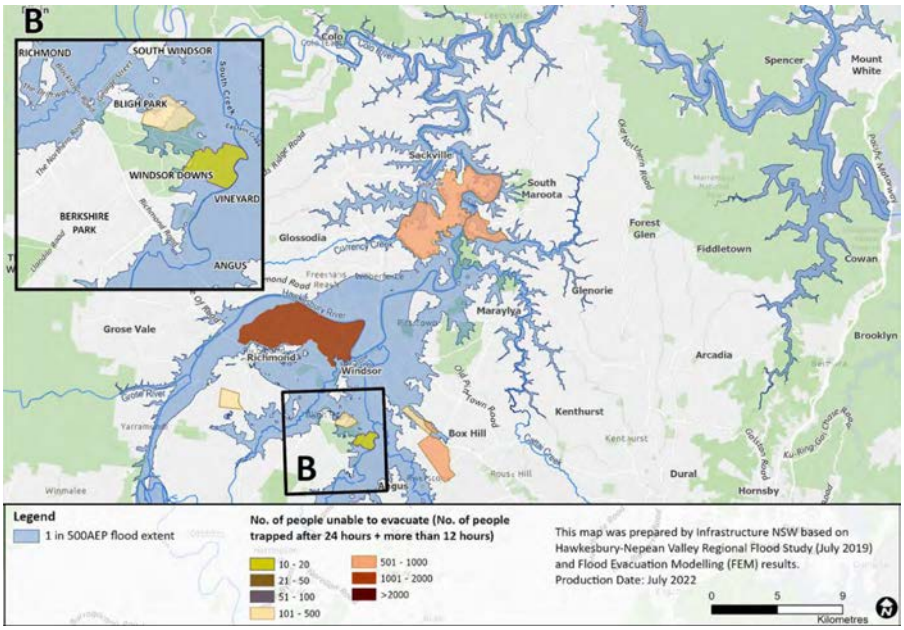


Figure 3. Extract of Figure 4.6 from NSW Government (2023) showing an estimated 6-fold increase in the number of people unable to evacuate in a 1 in 500 chance per year flood, under a medium mid-century climate change scenario with additional committed housing developments up to 2041, compared to the 2018 base case.

making its recommendation, the committee noted that the plan "does not demonstrate that the threat of bushfire risk to people, property and infrastructure can be appropriately mitigated [. . .] because: a. the technical information submitted in support [. . .], including traffic, evacuation and bushfire analysis and modelling, does not include adequate or accurate details to demonstrate that the risk of impacts from a bushfire, including safe evacuation in a bushfire emergency, is acceptable;" The case is significant because evacuation modelling played an essential role in the assessment of risk to a proposed future community from bushfires. It highlights the requirement for evacuation modelling to be "adequate and accurate" to be considered suitable for decision making. It emphasises the requirement for a tight integration between the fire progression modelling and evacuation modelling. Importantly, it demonstrates the demand for this capability for town planning in bushfire-prone areas.

An example of detailed evacuation modelling in Australia that was considered adequate and accurate for planning purposes, in this case by the State of New South Wales, including Infrastructure NSW, Transport for NSW, the NSW State Emergency Service, and the NSW Department of Planning and Environment, is the Flood Evacuation Model (FEM) developed by the CSIRO in collaboration with industry and academic partners (NSW Government, 2023). FEM was built to inform flood risk management planning in the Hawkesbury-Nepean river valley near Sydney, for use by government agencies and local councils responsible for integrated land use, transport and emergency planning in the area. FEM simulates the NSW State Emergency Service evacuation arrangements in current and future planned development scenarios (2018, 2026, and 2041) across a suite of representative flood events (from a 1-in-50 to 1-in-5000 chance per year). Through thorough analysis of results from a large ensemble of simulation runs, FEM provides an understanding of road network performance and capacity constraints for estimated evacuating traffic; assesses risk to life from saturated capacities and congested traffic flows against impending flooding; and estimates risk reduction from potential road upgrades (see Figure 3).

The Risk Analyst

Fire progression models like Phoenix, Aurora/Australis, and Spark are used within Australian emergency services (Faggian et al., 2017) at various levels of operation and time scales (see Table 1) from readiness to incident response. Personnel trained in the role of a fire behaviour analyst (FBAN) are responsible for collecting weather and fuel load information, simulating and studying fire progression under known and forecast conditions, and providing predictive insights to positioned or operational crews, towards reducing risk to lives, livelihoods, properties, and infrastructure. On days when fire danger is rated High, Extreme, or Catastrophic, an FBAN may run several fire progression simulations at the start of the day, prepare a summary on likely progression of select fires that had the potential to cause significant damage to communities and assets were they to ignite, and distribute the predictive insights to relevant teams. In ongoing fires, FBANs are installed in Incident management Teams and operate in

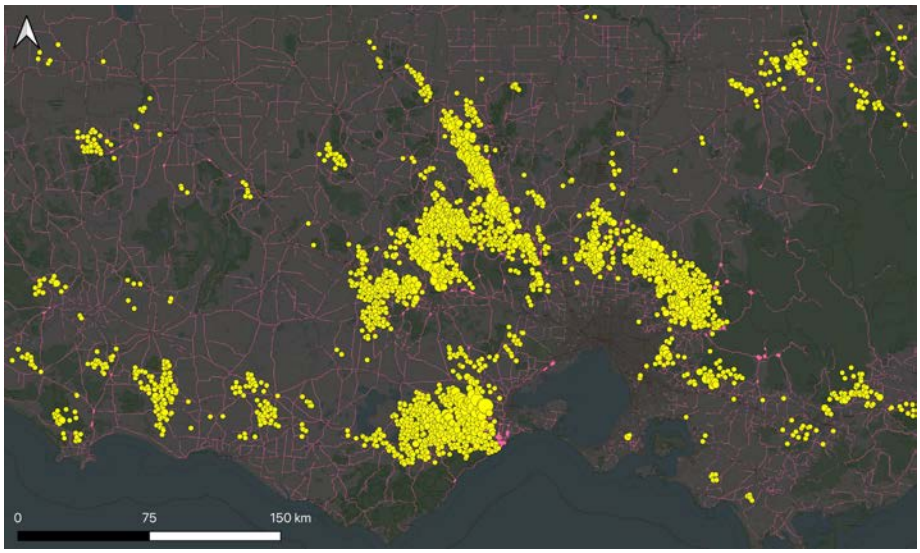


Figure 4. Example SAFER outputs showing evacuation hotspots in Victoria, Australia, as reported by Singh et al. (2023). Larger fire ignitions (yellow) are more problematic, while thicker road segments (pink) are more critical, for evacuating traffic.

Incident Control Centres at local, regional and state levels. Outside of Australia, the FBAN role is also employed in the USA. It is likely, under current emergency management arrangements, that the responsibility of running evacuation scenarios using outputs of fire progression models as inputs would fall initially on the FBAN. This highlights several important requirements for evacuation models: of timely, trusted, easily interpreted, explainable, and unambiguous outputs for accountable decision-making downstream.

Strategic long-term fire risk assessment is performed for large jurisdictions across a set of fire ignitions under different weather and wind conditions. Predictive Services within the Department of Energy, Environment, and Climate Action in Victoria, Australia, used "millions of fire simulations across the state to identify high risk areas, test different fire management scenarios and quantify risk reduction from delivered treatments" (Melero Blanca & Brien, 2023). Similarly, the Western Australian Department of Fire and Emergency Services used six million fire simulations to assess power network related fire risk from ignitions sparked at any power pole in the state's southwest (Parker & Allen, 2019). The CSIRO reported on the SAFER model, that consumes these very large sets of fire simulations to compute aggregate evacuation risk "hotspots" (see Figure 4 for Australian states (Singh et al., 2023). Noting the sheer size of the task, authors emphasise additional requirements for evacuation modelling capability in this context as: "The design goals for SAFER are (i) *scale*: supporting risk assessment for areas as large as Australian states and territories, and across hundreds of thousands of simulated fire ignitions, (ii) *speed*: computing state-wide evacuation risk metrics within hours, and (iii) *automation*: producing outputs automatically with no manual steps."

The Incident Controller and Response Crew

During a rapid-onset bushfire event, the Incident Management Team consisting of personnel representing all relevant emergency agencies is immediately activated in Australia, with explicit processes and protocols for emergency management and escalation. Within this team, it is the role of the Incident Controller (IC) to make the decision to evacuate or not, often under time pressure and with limited situation awareness. This is a decision that comes with enormous responsibility because if made incorrectly it can put people directly in harm's way.

Evacuation modelling capability that could provide fast and valuable insights into possible consequences for the community at risk, that the IC could take into account when considering the options, would be beneficial to the IMT as an additional source of assessment of the situation. However, for this to work, the technology must be trusted by the IC, allow exploration of emergency response scenarios the IC is considering in the situation, provide information on worst and likely consequences in forms that can be quickly understood by the IC, answer questions according to an explicit decision workflow the IC must follow, and highlight decision aspects that community safety outcomes are most sensitive to and may require careful attention. The reality is that the IC makes the decision to evacuate or not, for every bushfire that may interact with communities, evacuation models at hand or not. However, were such capability available, it would provide a valuable data-driven perspective to inform decisions. While this



Figure 5. Evacuation modelling use in Incident Management Team training as reported by Singh et al. (2022).

example is specific to one region and hazard, comparable emergency decision arrangements exist in many parts of the world and in the end a human is in charge of making the important decision on evacuation. The evacuation modelling capability must fundamentally cater to this user's requirements.

Singh et al. (2022) reported on the use of bushfire evacuation modelling in a multi-agency joint Incident Management Team training exercise organised by the Department of Environment, Land, Water and Planning at the dedicated Geelong Incident Control Centre in Victoria, in preparation for the 2021-22 bushfire season. During the live real-time training, evacuation modelling against the hypothetical fire was presented to the IMT in the form of analysis from several "what-if" evacuation scenarios of emergency response (see Figure 5). These scenarios included varying of timing and severity of warnings, partial or full implementation of manned traffic management points at intersections as per the local traffic management plan for emergencies, and combinations thereof. Outputs included map-based visualisations, videos of simulated evacuations against the fire progression, summaries of estimated consequences from each, and recommendations for the IC. Post-event reflection showed that evacuation modelling was a useful source of additional information for the IC who consulted the modelling on several occasions through the day. Authors noted that videos of evacuating vehicles were referenced by participants more than zone-based visualisations of evacuation metrics, hinting at the need for outputs that are fit-for-purpose. Post-event debriefing also identified areas of improvements including the ability to re-run scenarios in real-time to estimate reduction in risk to lives following IMT decisions.

EVACUATION MODELLING REQUIREMENTS FROM END USER PERSPECTIVES

We now return to an earlier point made in the Introduction that the adoption of evacuation models in disaster risk management requires addressing not just scientific or technological challenges, but societal barriers as well. In other words, the development of appropriate technological solutions must be guided by clearly articulated societal needs. Related to this point is the fact that, as is apparent from the previous section, different end users may have very different needs, and therefore it is imperative that modellers resist the temptation of a one-size-fit-all approach. Rather, each use case may call for a different approach, not just in the modelling techniques used, but also in the design of the corresponding software solution. At this point, an argument may be made that the term end users should be interpreted broadly, that is any individual or organization utilizing the outputs of a model in their decision making process. This does not necessarily imply that end users are responsible for executing the simulation scenarios themselves, but rather that the outputs of such scenarios should be made accessible to them.

As highlighted in the Introduction, the design process is inherently iterative, involving developers, researchers and end users. The objective during this phase should be to clarify the intended use of the model as much as possible, in order to create a solution that is fit for purpose. Relevant questions should include for example: What questions does the model need to be able to answer? What input data is required, and is it readily available? How fast does the model need to run? What kind of outputs does it need to produce, and in what form? Who is responsible for

running the models and producing results? If there is technology transfer, how is the risk of mishandling complex scientific software managed?

Irrespective of the specific use case considered, generally speaking a model of evacuation will need to address broad categories of key requirements, including:

- *Ability to set up multiple simulations scenarios*. This requirement includes the need for the model to easily ingest input data with minimal human intervention to miminimize time-consuming tasks, human errors, and enhance provenance processes and scenario reproducibility. An additional challenge specific to the use of these models in a response context is the ability to access live data, which may be challenging due to privacy reasons.
- *Capture of Important Evacuation Scenario Features*. The model should capture crucial features of the evacuation scenario under consideration. For instance, human behaviour modelling is critical for regional-scale evacuations but may be mostly irrelevant when identifying potential hotspots for evacuation risk at a larger scale (e.g. state-wide in Australia).
- *Reliability of models and outputs*. The issue of model validation for evacuation applications is an open question, as will be briefly discussed in the following section. Even if a full direct validation exercise may prove difficult to carry out, it is essential to consider quantification and adequate communication of uncertainty in the model outputs to prevent misrepresentation of scientific results with potentially devastating consequences.
- *Models as decision support systems*. Ultimately, an evacuation modelling tool is a decision support system, designed to assist end users in their decision-making process by providing them with the best available modelling data and is not intended as a replacement for human decisions. This implies that model outputs must be presented in a manner that facilitates a clear interpretation of results and minimizes the risk of judgment errors. For an IC for example, this means being able to easily compare the outputs of different simulation scenarios in a scientifically robust manner, to quickly gain an understanding of which plan may be the most effective under the conditions considered, taking into account any uncertainties and unknowns factors.

We argue that the process of identifying these requirements necessitates an extensive consultative approach well before any software development begins. Throughout this process, all parties involved should develop three essential design documents: a user requirements document, which lists all requirements that the technology needs to address, a risk register outlining roles, responsibilities, and risk mitigation strategies, and a strategic roadmap defining the development path with a clearly defined end goal. Finally, these documents should be revisited as the developments progress, with ongoing testing and feedback provided by end users. Only through such a thorough consultative process can we bridge the current gap in the use of computational modelling for evacuation management.

A BRIEF REVIEW OF MODELLING LARGE-SCALE EVACUATION

As the number of evacuation models continues to grow, user experiences and needs are pivotal for broader adoption, spanning community, government, and industry. Lovreglio et al., 2020 conducted a comprehensive international online survey in 2018, primarily focusing on pedestrian and indoor evacuation models but with implications for traffic evacuation. The survey highlighted key factors influencing model selection, emphasising validation and verification (V&V), data output, model documentation, continual development, and software usability. Notably, V&V emerged as a significant concern, reflecting users' demand for trustworthy model outputs. However, applying V&V to large-scale outdoor scenarios, especially in the context of wildfires, proves challenging due to the less constrained spatial and temporal behaviour compared to structured evacuations from buildings. While existing V&V documents provide guidelines for specific contexts, the lack of universally applicable procedures remains a challenge. Efforts by organisations like the National Institute of Standards and Technology (NIST) (Ronchi et al., 2013, 2016) and the International Standardisation Organisation (ISO) (ISO, 2020) are progressing toward standardised protocols, but challenges persist in defining comprehensive and universally applicable V&V processes for evacuation models. Ronchi et al., 2023 contribute to this with a verification protocol for wildfire evacuation models, acknowledging the difficulties and proposing tests for accurate implementation of modeling layers. It is important to note that the survey work by Lovreglio et al., 2020 is highly valuable for indoor fires but may require adaptation for large-scale evacuations. Additionally, the survey's utility is reinforced by the inclusion of respondents from diverse backgrounds, including end users outside of academia, accounting for 234 participants

from 41 countries and various professional domains. However, despite the insights provided by the survey, none of the covered models, including the pedestrian and building evacuation models, are currently being used for strategic planning, readiness, or response, underlining the ongoing challenges in applying these tools to real-world scenarios beyond academic and research contexts.

Notwithstanding, there are several simulation platforms for evacuation currently available. WUI-NITY, a modular simulation platform built on the Unity 3D game engine, effectively integrates fire, pedestrian, and traffic models to forecast evacuation behaviour during bushfire events (Kuligowski et al., 2022). This freely available platform allows for the simulation and visualisation of human behaviour, bushfire development, and evacuation scenarios. WUI-NITY's capability to import data from external wildfire models, such as FARSITE (Finney, 1998), and its implementation of a traffic model based on the Lighthill-Whitham-Richards (LHR) model (Li et al., 2012) contribute to its versatility. Recent studies, including a community drill in Roxborough, Colorado, and simulations of California wildfires, have validated WUI-NITY's performance, highlighting its potential for planning and real-time decision-making applications with its simplified yet effective models (Gwynne et al., 2023). The platform's modularity allows for the coupling of different sub-models, and its accessibility to a wide audience makes it a valuable tool for future model calibration and validation efforts.

There are other simulation platforms for evacuation currently available. SEEKER is a simulation platform integrated with an archetype-based behaviour model and employed by Victorian emergency services to understand community responses in bushfire evacuation scenarios (Marquez et al., 2022; Singh & Padgham, 2017; Singh et al., 2021). SEEKER takes as input fire progression simulation outputs from model like Phoenix and Spark (Faggian et al., 2017), a routable representation of the road network from Open Street Map, and a synthetic representation of individuals in the impacted area, and simulates community responses, considering various emergency response levers like warning messages and traffic control. Incorporating situation-based cognitive decision-making and MATSim traffic simulation, the platform provides detailed outputs on vehicles' trajectories and individuals' decisions during evacuations, including any intermediate trips to locations other than their final destinations. These outputs aid in identifying congestion bottlenecks, understanding diverse evacuation outcomes against a range of behaviour assumptions, and informing planning and preparedness. Simulation analysis supports operational decision-making, and the platform allows virtual testing of response strategies, including emergency messages and traffic control measures.

EXODUS is a suite of software tools designed for simulating the evacuation of large populations from various enclosures, including buildings, aircraft, and ships (Gwynne et al., 2006). Its core components include the Occupant, Movement, Behaviour, Toxicity, and Hazard sub-models (Galea & Galparsoro, 1994; Galea et al., 2000). The software employs rule-based descriptions, utilising heuristics to determine the progressive motion and behaviour of individuals. The Occupant sub-model in EXODUS characterises the response of individuals during evacuations based on their personal attributes (i.e., physical, positional, personal and hazard effects) influencing the individual's decision-making process and enabling the simulation to track their progress and behaviour during the evacuation process (Owen et al., 1996; Zhao et al., 2017). The Movement sub-model governs the travel speed of individuals during evacuations, taking into account factors such as terrain, obstacles, and queues (Gwynne et al., 2019). It assigns different travel speeds, such as run, walk, leap, and crawl, based on the specific conditions each evacuee encounters during the evacuation process. The Behaviour sub-model, influences occupants' responses and evacuation dynamics based on their attributes and environmental conditions (Galea et al., 1996). The Toxicity sub-model assesses the physiological impact of the environment on occupants, considering hazards like elevated temperature, thermal radiation, and toxic gases (Gwynne et al., 2001). The Hazard sub-model distributes hazards over time and location. EXODUS is well-suited for diverse applications, including planning, preparedness, readiness, response, emergency training, and community education. It provides valuable insights into evacuation scenarios and includes data analysis tools for effective result interpretation, as has been demonstrated in various publications (e.g., Galea, 1998; Gwynne et al., 2000).

DISCUSSION

Evacuation modelling is most useful for emergency management and planning when it supports the myriad of concerns of all stakeholders and incorporates realistic human behaviours and actions over idealised or optimised ones. Adoption of evacuation modelling capability for decision support involves a process of co-learning and co-production between emergency management stakeholders and technology providers. For emergency agencies to use evacuation models to inform decisions that impact lives, an often overlooked yet essential ingredient is time: required to understand, explore, adapt, and accept the modelling capability confidently into decision workflows. Model providers equally require time to shift evacuation modelling from a mathematical concept to a human-centred one.

The role of FBANs along with fire progression models in the emergency services is formalised and has existed for over a decade. FBANs are regarded with a high level of trust within teams, and that trust extends to the timeliness and quality of FBAN outputs (Begg et al., 2021). No such role exists in the predictive services for evacuation related modelling capability yet. This is partly because the need for evacuation planning is relatively new in Australia (Binskin et al., 2020; Teague et al., 2010) and current state-of-the-art evacuation modelling capability sits around Technology Readiness Level (TRL) 6 (technology demonstrated in a relevant environment), e.g., SEEKER (Singh et al., 2022), compared to fire progression models that have reached TRL9 (full system proven in operational environment), e.g., SPARK (Bally, 2023). Outside of technology readiness, other determinants of adoption of technology in emergency services include individual (e.g., perceived usefulness and ease of use, individual experience in emergency management), organisational (e.g., organisations IT infrastructure, interoperability and readiness, financial resources, decision-making processes), and contextual determinants (e.g., social norms and culture, and IT fitness for purpose) (Gulatee et al., 2020). Overcoming individual, organisational, and contextual challenges takes time, particularly in emergency management where reliance on technology for decisions that directly impact lives comes from trust, and the progression of technology over the final rungs to TRL9 is a necessarily slow process, and rightly so.

The Community-Based Bushfire Management program (Macken, 2019) is a successful example of a government funded program supporting emergency services and communities working together in Australia, in a community-led partnership to appreciate and plan for bushfire risk, understand roles and responsibilities, increase awareness and education, build trusting relationships, and ultimately improve safety in bushfires. Use of evacuation modelling to support these kinds of collaborative initiatives (see examples is Figure 2 and Dennis (2022)) shows promise in reaching a shared understanding and improving community resilience.

Compared with tools for modelling natural hazards like wildfires and floods, technology for modelling and simulating human behaviour during such hazards is relatively new. However, modern day computing infrastructure is rapidly enabling the use of this technology in tandem with natural systems modelling. We presented examples of the ways in which this technology is starting to be considered, beyond its obvious use in planning and preparedness, in also training and community education. What is unique here is a feedback loop that could positively alter community outcomes during natural hazards: since simulations of community and emergency responses today can help inform resilient behaviours of tomorrow.

REFERENCES

Atkin, G. (2023). Blackwood Emergency Planning Expo. *The Blackwood Times*, *15*(16.3). https://theblackwoodtimes.com.au/wp-content/uploads/2023/10/2 BWT-OctNov023-Colour-WEBsm.pdf

Bally, J. (2023). Spark Operational: Development to implementation of Australia's next generation bushfire behaviour simulator. https://www.afac.com.au/docs/default-source/2023-Conference-proceedings/416-spark-operational-development-to-implementation-of-australia-39-s-next-generation-bushfire-behaviour-simulator john-bally.pdf

Begg, C., Dwyer, G., Neale, T., & Pollock, I. (2021). *Established and emerging uses of predictive services in Victoria*. Bushfire; Natural Hazards CRC. https://www.bnhcrc.com.au/sites/default/files/managed/downloads/established_and_emerging_uses_of_predictive_services_black_summer_final_report.pdf

Beverley, J. (2023). Rethinking Risk – Some Insights from Wildfires in Canada (Keynote Presentation). *AFAC Conference*. https://www.afac.com.au/insight/operations/article/current/keynote-rethinking-risk-some-insights-from-wildfires-in-canada

Beverly, J. L., & Bothwell, P. (2011). Wildfire evacuations in Canada 1980–2007. *Natural Hazards*, *59*(1), 571–596. https://doi.org/10.1007/s11069-011-9777-9

Binskin, M., Bennett, A., & Macintosh, A. (2020). Royal Commission into National Natural Disaster Arrangements Report Chapter 12: Evacuation planning and shelters. *Canberra: Commonwealth of Australia*.

Borchers Arriagada, N., Palmer, A. J., Bowman, D. M., Morgan, G. G., Jalaludin, B. B., & Johnston, F. H. (2020). Unprecedented smoke-related health burden associated with the 2019–20 bushfires in eastern Australia. *Medical Journal of Australia*, *213*(6), 282–283. https://doi.org/10.5694/mja2.50545

Bulumulla, C., Padgham, L., Singh, D., & Chan, J. (2017). The Importance of Modelling Realistic Human Behaviour When Planning Evacuation Schedules. *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, 446–454. Retrieved October 16, 2018, from http://dl.acm.org/citation.cfm?id=3091125.3091192

Chen, J., Shi, T., & Li, N. (2021). Pedestrian evacuation simulation in indoor emergency situations: Approaches, models and tools. *Safety Science*, *142*, 105378.

Community Outcomes and Recovery Subcommittee (CORS). (2022, October). *Australian Disaster Recovery Framework Version 3.0*. Australia-New Zealand Emergency Management Committee (ANZEMC).

CSIRO Evacuation Modelling Group. (2023, December). Creating Evacuation Decision Support Tools for Australia. Retrieved December 19, 2023, from https://research.csiro.au/evacuation/

Dennis, L. (2022, July). Evacuation Mapping Workshop to offer insights into bushfire management at Chewton. Retrieved December 18, 2023, from https://castlemainemail.com.au/latest-news/2022/07/29/evacuation-mapping-workshop-to-offer-insights-into-bushfire-management-at-chewton/

Faggian, N., Bridge, C., Fox-Hughes, P., Jolly, C., Jacobs, H., Ebert, B., & Bally, J. (2017, June). *Bushfire Predictive Services Final Report: An evaluation of fire spread simulators used in Australia*. Bureau of Meteorology.

Finney, M. A. (1998). *FARSITE, fire area simulator-model development and evaluation*. US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Galea, E. R. (1998). A general approach to validating evacuation models with an application to EXODUS. *Journal of Fire Sciences*, *16*(6), 414–436.

Galea, E. R., & Galparsoro, J. P. (1994). A computer-based simulation model for the prediction of evacuation from mass-transport vehicles. *Fire Safety Journal*, *22*(4), 341–366.

Galea, E. R., Gwynne, S. M., Lawrence, P., Filippids, L., BlackShields, D., & Cooney, D. (2000). BuildingEXODUS User Manual, Version 3.01.

Galea, E. R., Owen, M., & Lawrence, P. J. (1996). Computer modelling of human behaviour in aircraft fire accidents. *Toxicology*, *115*(1-3), 63–78.

Gamage, P. (2022, October). Showcase of evacuation decision support tools at Castlemaine Emergency Response Expo 2022. Retrieved December 18, 2023, from https://research.csiro.au/evacuation/evacuation-support-tools/

Gamage, P. (2023, October). A day well spent at the Korweinguboora expo talking all things evacuation with the community. Retrieved December 18, 2023, from https://research.csiro.au/evacuation/korweinguboora-expo-2023/

Goto, Y., Affan, M., Nurdin, Y., Yuliana, D. K., et al. (2012). Tsunami evacuation simulation for disaster education and city planning. *Journal of Disaster Research*, *7*(1), 92–101.

Government of Western Australia. (2023, December). North Stoneville Structure Plan 34. Retrieved December 14, 2023, from https://www.wa.gov.au/government/document-collections/north-stoneville-structure-plan-34

Gulatee, Y., Yuan, Q., Gasco-Hernandez, M., Gil-Garcia, J. R., Sutherland-Mitzner, M., & Pardo, T. A. (2020). Technology adoption for emergency preparedness and response in rural areas: Identifying the main determinants. *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance*, 469–476. https://doi.org/10.1145/3428502.3428574

Gwynne, S. M., Galea, E. R., & Lawrence, P. J. (2006). The introduction of social adaptation within evacuation modelling. *Fire and Materials: An International Journal*, *30*(4), 285–309.

Gwynne, S. M., Galea, E. R., Lawrence, P. J., & Filippidis, L. (2001). Modelling occupant interaction with fire conditions using the BuildingEXODUS evacuation model. *Fire Safety Journal*, *36*(4), 327–357.

Gwynne, S. M., Galea, E. R., Lawrence, P., Owen, M., & Filippidis, L. (2000). Adaptive decision-making in BuildingEXODUS in response to exit congestion. *Fire Safety Science*, *6*(2000), 1041–1052.

Gwynne, S. M., Owen, M., Galea, E. R., Filippidis, L., & Lawrence, P. (2019). Adaptive decision-making in response to crowd formations in BuildingEXODUS. In *Evacuation from fires* (pp. 73–96). Routledge.

Gwynne, S. M., Ronchi, E., Wahlqvist, J., Cuesta, A., Gonzalez Villa, J., Kuligowski, E. D., Kimball, A., Rein, G., Kinateder, M., Benichou, N., et al. (2023). Roxborough Park community wildfire evacuation drill: Data collection and model benchmarking. *Fire Technology*, *59*(2), 879–901.

ISO. (2020). Fire safety engineering—verification and validation protocol for building fire evacuation models ISO/DIS 20414.

Keelty, M. (2011). *A Shared Responsibility-the report of the Perth Hills Bushfire February 2011 review*. Government of Western Australia.

Kuligowski, E. (2021). Evacuation decision-making and behavior in wildfires: Past research, current challenges and a future research agenda. *Fire Safety Journal*, *120*, 103129. https://doi.org/10.1016/j.firesaf.2020.103129

Kuligowski, E., Ronchi, E., Wahlqvist, J., Gwynne, S. M., Kinateder, M., Rein, G., Mitchell, H., Benichou, N., & Kimball, A. (2022). Evacuation modelling for bushfire: The WUI-NITY simulation platform. *Australian Journal of Emergency Management*, *37*(4), 40–43. https://knowledge.aidr.org.au/media/9660/ajem-13-2022-04.pdf

Li, J., Chen, Q.-Y., Wang, H., & Ni, D. (2012). Analysis of LWR model with fundamental diagram subject to uncertainties. *Transportmetrica*, *8*(6), 387–405.

Lovreglio, R., Ronchi, E., & Kinsey, M. J. (2020). An online survey of pedestrian evacuation model usage and users. *Fire Technology*, *56*(3), 1133–1153.

Macken, F. (2019). Community-based bushfire management in Victoria. *Australian Journal of Emergency Management, The*, *34*(4), 9–10. https://knowledge.aidr.org.au/media/7271/ajem 102019-04.pdf

Marquez, L., Gamage, P., Singh, D., Lemiale, V., Dess, T., Ryan, L., & Ashton, P. (2022). SEEKER: A Web-Based Simulation Tool for Planning Community Evacuations. *Proceedings of the 2022 ISCRAM Asia Pacific Conference*. https://idl.iscram.org/files/leoreymarquez/2023/2476_ LeoreyMarquez etal2023.pdf

McLennan, J., Ryan, B., Bearman, C., & Toh, K. (2019). Should We Leave Now? Behavioral Factors in Evacuation Under Wildfire Threat. *Fire Technology*, *55*(2), 487–516. https://doi.org/10.1007/s10694-018-0753-8

Melero Blanca, E., & Brien, P. (2023). Using fire science and bushfire risk modelling to support end-to-end, risk-based bushfire management in Victoria. https://www.afac.com.au/docs/default- source/2023-Conference-proceedings/259-using-fire-science-and-bushfire-risk-modelling-to-support-end-to-end-risk-based-bushfire-management-in-victoria estrella-melero-blanca.pdf

NSW Government. (2023, May). *Hawkesbury-Nepean Valley flood evacuation modelling to inform flood risk management planning*. https://www.nsw.gov.au/sites/default/files/2023-08/Flood%20Evacuation%20Modelling%20Report%20May%202023.pdf

NSW Rural Fire Service. (2023, December). Details on the annual public awareness campaign conducted by the NSW RFS. Retrieved December 14, 2023, from https://www.rfs.nsw.gov.au/news-and-media/advertising

Owen, M., Galea, E. R., & Lawrence, P. (1996). *Advanced occupant behavioural features of the Building-EXODUS evacuation model* (Vol. 9). CMS Press.

Parker, J., & Allen, A. (2019). Utilisation of fire spread simulators to assess power network fire risk. *Australian Journal of Emergency Management, Monograph No. 5*, 28–34. https://www.aidr.org.au/media/7379/monograph-no5-extended-abstracts-final.pdf

Perth Hills Action Group. (2023, December). Save Perth Hills. Retrieved December 14, 2023, from https://www.saveperthhills.net

Robertson, J., Singh, D., Marquez, L., Lemiale, V., Ashton, P., Dess, T., Halliday, J., Gamage, P., & Prakash, M. (2021). Accounting for seasonal populations in bushfire evacuation modelling and planning: A Surf Coast Shire case study. *AFAC Conference*. https://www.afac.com.au/events/proceedings/05-10-21/article/accounting-for-seasonal-populations-in-bushfire-evacuation-modelling-and-planning-a-surf-coast-shire-case-study

Ronchi, E., Wahlqvist, J., Ardinge, A., Rohaert, A., Gwynne, S. M., Rein, G., Mitchell, H., Kalogeropoulos, N., Kinateder, M., Be´nichou, N., et al. (2023). The verification of wildland-urban interface fire evacuation models. *Natural Hazards*, *117*(2), 1493–1519.

Ronchi, E., Gwynne, S. M., Rein, G., Wadhwani, R., Intini, P., & Bergstedt, A. (2017). *E-Sanctuary: Open multi-physics framework for modelling wildfire urban evacuation*. Fire Protection Research Foundation Quincy.

Ronchi, E., Kuligowski, E. D., Nilsson, D., Peacock, R. D., & Reneke, P. A. (2016). Assessing the verification and validation of building fire evacuation models. *Fire Technology*, *52*, 197–219.

Ronchi, E., Kuligowski, E. D., Reneke, P. A., Peacock, R. D., & Nilsson, D. (2013). The process of verification and validation of building fire evacuation models.

Singh, D., Ashton, P., Dess, T., Harper, M., Kuligowski, E., Gamage, P., Marquez, L., Lemiale, V., Halliday, J., McKenzie, R., & Prakash, M. (2022). Bushfire evacuation decision support system use in incident management training. *Australian Journal of Emergency Management*, *37*(4), 73–76. https://knowledge.aidr.org.au/resources/ajem-october-2022-bushfire-evacuation-decision-support-system-use-in-incident-management-training/

Singh, D., & Padgham, L. (2017). Emergency Evacuation Simulator (EES)- A tool for planning community evacuations in Australia. *IJCAI*, *17*, 5249–5251.

Singh, D., Strahan, K., McLennan, J., Robertson, J., & Wickramasinghe, B. N. (2021). What will they do? Modelling self-evacuation archetypes. *CoRR*, *abs/2105.12366*. https://arxiv.org/abs/2105.12366

Singh, D., Wirian, S., Taylor, S., Gamage, P., Marquez, L., & Lemiale, V. (2023). SAFER: A tool to compute bushfire evacuation hotspots at scale. https://www.afac.com.au/insight/risk/article/current/366.-safer-a-tool-to-compute-bushfire-evacuation-hotspots-at-scale

Smith, R., Parker, J., Choney, G., Moss, J., & Lia, J. (2015). *Report On - Bushfire Behaviour and house damage and destruction during the Pakerville, Mt Helena & Stoneville Bushfire on the 12 January 2014*. Government of Western Australia.

Strahan, K., Whittaker, J., & Handmer, J. (2018). Self-evacuation archetypes in Australian bushfire. *International Journal of Disaster Risk Reduction*, *27*, 307–316. https://doi.org/10.1016/j.ijdrr.2017.10.016

Strahan, K. W., Whittaker, J., & Handmer, J. (2019). Predicting self-evacuation in Australian bushfire. *Environmental Hazards*, *18*(2), 146–172. https://doi.org/10.1080/17477891.2018.1512468

Teague, B., McLeod, R., & Pascoe, S. (Eds.). (2010, July). *2009 Victorian Bushfires Royal Commission Final Report*. Parliament of Victoria, 2009 Victorian Bushfires Royal Commission.

Wahlqvist, J., Ronchi, E., Gwynne, S. M., Kinateder, M., Rein, G., Mitchell, H., Be´nichou, N., Ma, C., Kimball, A., & Kuligowski, E. (2021). The simulation of wildland-urban interface fire evacuation: The WUI-NITY platform. *Safety Science*, *136*, 105145.

Whittaker, J. (2019). Ten years after the black Saturday fires, what have we learnt from post-fire research? *Australian Journal of Emergency Management, The*, *34*(2), 32–37.

Wong, S. D., & Broader, J. C. (2020). Review of California Wildfire Evacuations from 2017 to 2019. https: //doi.org/10.7922/G29G5K2R

Zhao, D., Zhang, Q., Liu, P., Sun, M., Wang, X., Zhang, X., & Wang, X. (2017). Simulation on occupant evacuation at a public site based on SMARTFIRE and building EXODUS. *Studies in Engineering and Technology*, *5*(1), 15.

Wildfire Evacuation Challenges in Multi-ethnic Communities: An Information Communication Perspective

Duo Meng
Beijing Normal University, China
duomeng@webmail.hzau.edu.cn

Hu Jun
Beijing Normal University, China
hujun0214@bnu.edu.cn

Chen Anying
Jinan University, China
chenay@jnu.edu.cn

Fang Zhetao
China People's Police University, China
1226442719@qq.com

ABSTRACT

The areas with high wildfire risk of wildfire are inhabited by multiple ethnic groups in China. In the event of a wildfire, there are many challenges for the government to release evacuation information to residents in multi - ethnic communities. Based on the Shannon-Weaver model of communication, a whole-process communication model of evacuation information is constructed to reflect the challenges during the whole process, including the dynamic variability, regional difference and cross-language expression of evacuation information at the stage of information generation, the inadequate communication capability of information carriers and instability of communication infrastructures at the stage of information transmission, and comprehension impairment and irrational perception at the stage of information perception. In response to these challenges, the authors also make some suggestions from the intelligent information technology, crisis response system and offline communication channels.

Keywords

Wildfire, evacuation, multi-ethnic, communication.

INTRODUCTION

Serious wildfires have occurred in many areas around the world recently. In addition to the result of ecological and environmental issues, such as declining air quality and vegetation destruction, they also pose a threat to human life and property safety. When a wildfire breaks out, except taking appropriate fire extinguishing, evacuating the residents around the fire in time is also important. The current research on wildfire evacuation is mainly carried out from the following aspects: (a) the influence factors on wildfire evacuation decisions. McCaffrey et al. (2018), McLennan et al. (2019), and Toledo et al. (2018) have analyzed the factors that affect residents' evacuation actions in wildfire scenarios. (b) the modeling and analysis of the evacuation process. Wang et al. (2021) constructed an evacuation model of forest tunnel fire based on the characteristics of smoke diffusion. Beloglazov et al. (2016) constructed a simulation of wildfire evacuation with dynamic factors and model composition, and evaluated the effectiveness of an evacuation. Cohn et al. (2006) analyzed the evacuation behavior during wildfires through case studies. (c) the evacuation risk from the perspective of local community and infrastructure. Cova et al. (2013) analyzed the limits of infrastructure on wildfire evacuation vulnerability, and Beverly et al. (2011) analyzed the impact of housing and roads during the wildfire evacuation process.

Some studies have also focused on the issue of information communication for wildfire evacuation. When a wildfire occurs, it is crucial for the public to receive timely information on evacuation (Stidham et al., 2011). McCaffrey et al. (2013) highlighted the importance of disseminating information in as timely a manner as possible

and through sources that allow as much interaction as possible with information providers. Li et al. (2021) demonstrated the efficacy of using social media to aid evacuations. Cao et al. (2016) believed that an optimal communication approach would be to couple map designs with several imperative textual descriptors, and Vélez et al. (2017) found that television is the primary source for information seeking during wildfires. Generally, these studies are conducted from the view of information receivers, namely the local residents at risk during a wildfire, and analyzed their attitudes and tendencies towards different information sources. However, there are few relevant research from the perspective of information communication process, and the information characteristics such as the timeliness and accuracy, which are also meaningful for wildfire evacuation, are not considered in-depth. Furthermore, community characteristics like multi-ethnic could also exert influence on the communication process, but relevant research is also lacked.

In this research, we will discuss the challenges of information communication during the wildfire. The study areas are the areas with high risk of forest fire in China, which are also the regions with multi-ethnic communities. Based on fire statistics, the current wildfire risk of China is presented. Then, from the perspective of evacuation information communication, this paper analyzes the challenges of information communication in these multi - ethnic areas with high wildfire risk. Finally, a whole-process communication model of evacuation information is constructed to reflect the challenges during the information communication process, and some suggestions are also provided.

THEORETICAL FOUNDATION

Getting the urgent information is the primary step before the public can take response actions. In general, whether the public can get the information in time is related to three factors: (1) The characteristics of the messages. The characteristics of the messages include the languages in which the messages are expressed and the forms of the messages (text, video, audio, etc.), which will have an impact on the accuracy, reliability and understandability of the information. (2) The transmission capacity of information carriers. For different information carriers, such as television, radio, online social media, and face-to-face communication, the speed and scale of information transmission are different, affecting the timeliness and population coverage of the information. (3) The comprehension capacity of the information receivers. The information needs to be understood and recognized after receiving it, and the comprehension capacity is related to the individual characteristics of the information receivers, including the age, career, education level, audio-visual condition, previous disaster experience and so on.

In this research, we mainly focus on the communication process of evacuation information during a wildfire. A classic theoretical model in communication research is the Shannon- Weaver Model (see Figure 1.), a process model of communication proposed by Claude Shannon and Warren Weaver in 1948 (Shannon, 1948). It is a transmission model consisting of six elements:

- Sender, the information source that produces a message;
- Encoder, the transmitter that encodes the message into signals;
- Channel, to which signals are adapted for transmission;
- Decoder, the reception that decodes (reconstructs) the message from the signal;
- Receiver, the destination where the message arrives.
- Noise, a dysfunctional factor.

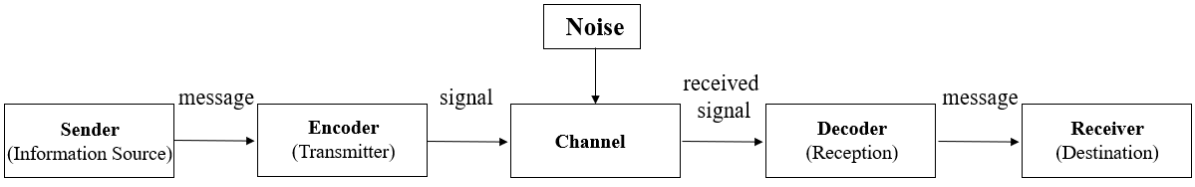


Figure 1. The Shannon-Weaver Model of Communication

The main feature of this model is the introduction of the concept of "noise", which refers to any interference with the message travelling along the channel and may lead to the signal received being different from that sent. It is a very linear model, and it does not even include a feedback loop. In our research, we focus on the linear process from official release to public reception of emergency information, so the Shannon-Weaver Model is adopted.

In this model, the sender will generate and send out the evacuation information to the receivers. For example, the local government or fire department can be considered as the sender during a wildfire, while the receivers are the

residents in the wildfire-affected regions. And the intermediate process from encoder to decoder can be regarded as the process of information transmission, including the channels of television, radio, SMS, online social media and so on. In addition, according to this theory, noise will occur in the process of information transmission, such as the unofficial and untrue rumors.

RESEARCH AREA

In this research, we will discuss the wildfire evacuation challenges in the regions with high risk of wildfire in China. There are great differences in natural conditions such as topography and climate among different regions in China, leading to the obvious regional differences in the risk of wildfire. Forest fire is a typical wildfire characterized by sudden and intense outbreaks, significant destruction, and challenging rescue efforts . The relevant data on forest fires in China from 2012 to 2021 provided by the *China Forestry Statistical Yearbooks* (2012-2021) reflects the occurrence of forest fires during this decade, and Figure 2 represents the total number of forest fires in each province of China according to the data. It can be observed that over the past decade, forest fires have predominantly occurred in the central and southern regions of the country, which is closely associated with differences in natural conditions such as topography and climate among these regions. In the north, the forest fire mainly occurs in Inner Mongolia. Specifically, Inner Mongolia, Shaanxi, Sichuan, Hubei, Hunan and Guangxi are identified as regions with a high incidence of forest fires and the total number of forest fires in Guangxi exceeded 4,000.

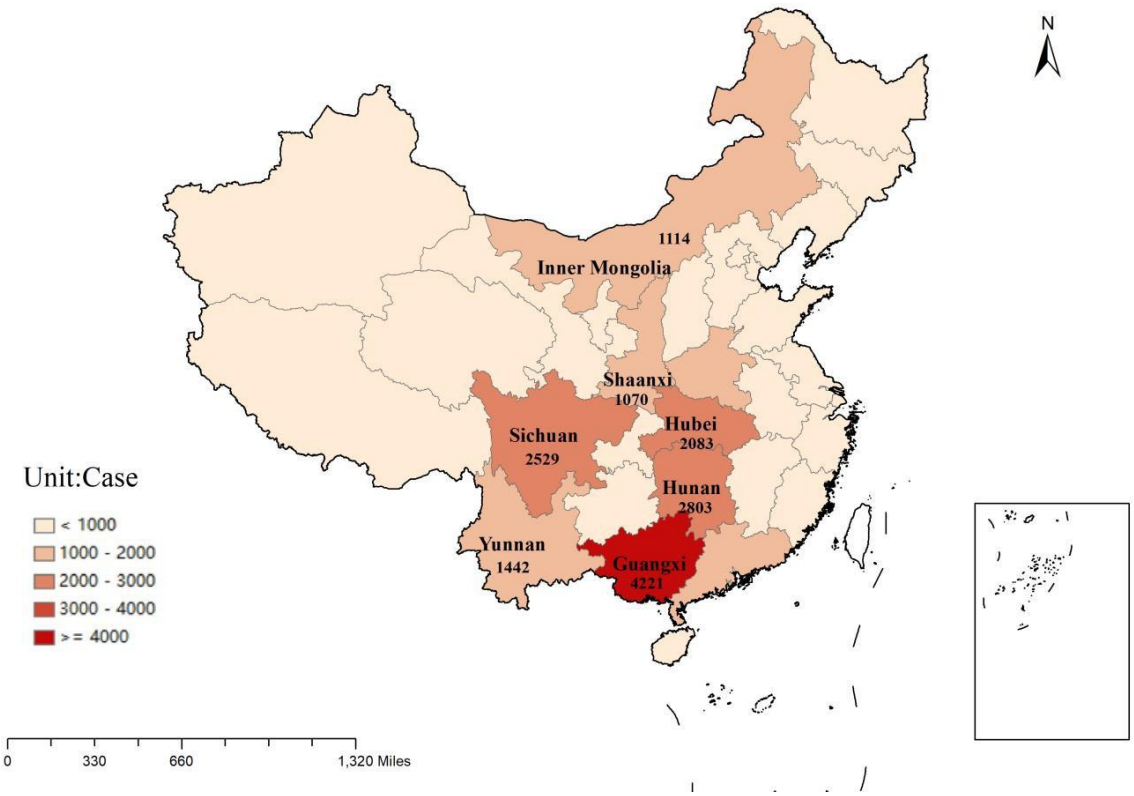


Figure 2. Total number of forest fires in each province of China (2012-2021)

The consequences of forest fires are serious. Figure 3 shows the number of forest fire casualties in China. Over the past decade, a total of 480 individuals have been reported to be injured or lost their lives due to forest fires, and the highest number of casualties was recorded in 2014, reaching 112 individuals. The provinces of Sichuan, Yunnan, Guangxi, Hunan, and Shaanxi have experienced a relatively higher number of casualties from forest fires. With regards to the damage to the ecology, the total damaged forest area has reached 133,000 hectares over the past decade. Figure 4 shows the affected forest area and total fire area in China. In 2014, the total fire area reached a peak of 55,341.4 hectares, and the affected forest area also reached a relatively high level of 19,111.1 hectares. In the subsequent years, both the affected forest area and total fire area showed a fluctuating decline and respectively reached 4,458 hectares and 14,125 hectares in 2021. Overall, the ecological damage caused by forest fires in China has been gradually decreasing, but the potential dangers still exist and should not be neglected.

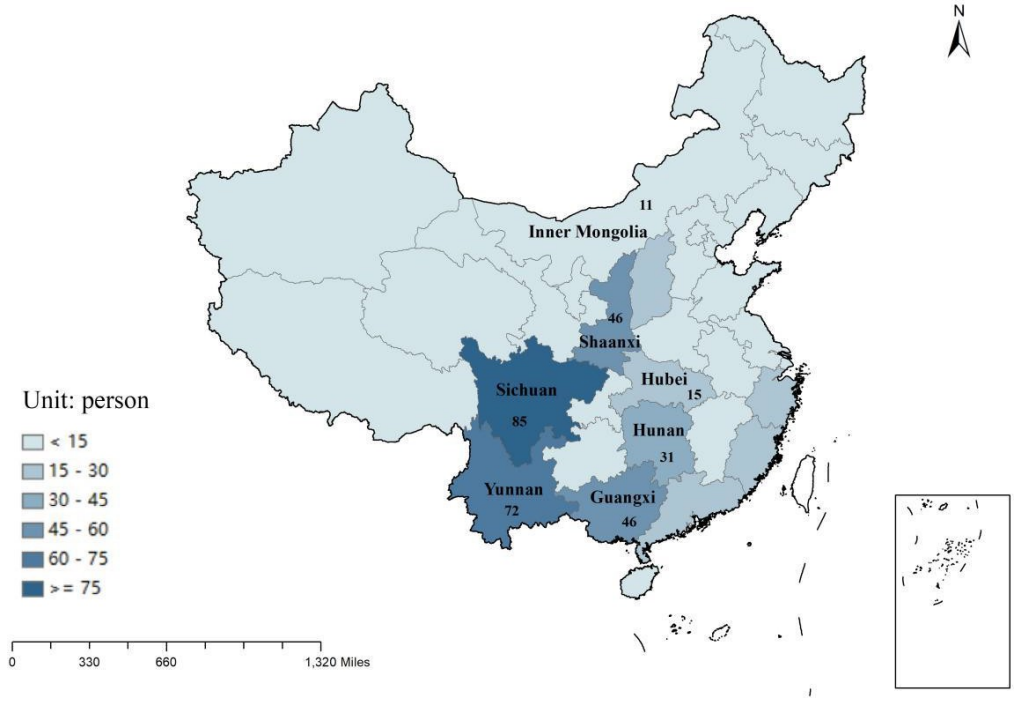


Figure 3. Total number of forest fire casualties in each province of China (2012-2021)

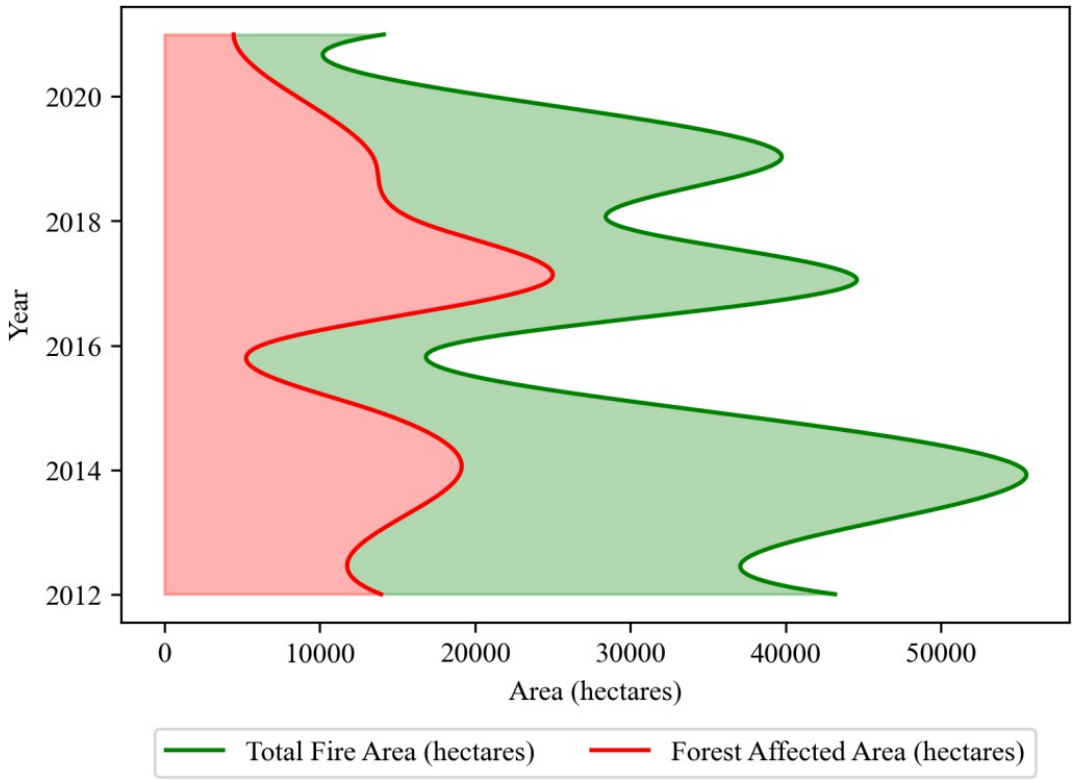


Figure 4. Total fire area and forest affected area over years (2012-2021)

Meanwhile, those serious and massive fire accidents should also be taken seriously, as they have caused serious casualties or losses. A serious case was happened in Liangshan, Sichuan Province, on March 30, 2019, which has caused 31 deaths. The brief information of several serious and massive large fire cases is shown in Table1.

Table 1. Cases of serious and massive forest fires happened in China (2012-2021)

Date	Location	Casualties or losses
Apr.16,2014	Kunming, Yunnan Province	210 hectares of forest burned
May.02,2017	Daxinganling, Inner Mongolia	About 8282.58 hectares of forest burned
Mar.30,2019	Liangshan, Sichuan Province	31 deaths
Dec.05,2019	Lianping, Guangdong Province	427.3 hectares of forest burned
Mar.30,2020	Liangshan, Sichuan Province	19 deaths and 3 wounded, 427.3 hectares of forest burned
Jun.07,2020	Changjiang, Hainan Province	About 358 hectares of forest burned

CHALLENGES: INFORMATION COMMUNICATION ON WILDFIRE EVACUATION

When a wildfire occurs, the timely information on evacuation is important to the public. The government will release evacuation information to residents in those communities which are deemed to be in need of evacuation to minimize potential casualties. However, there are many challenges in the actual implementation of the program.

Multi-ethnic Characteristics of Information Receivers

In the above section, we have analyzed the current situation of forest fire risk in China, and the areas with high wildfire risk include Inner Mongolia, Shaanxi, Sichuan, Hubei, Hunan, Yunnan and Guangxi provinces. In fact, these areas are also ethnically diverse. Figure 5 illustrates the current status of multicultural integration in these provinces according to *National Bureau of Statistics* (Office of the Leading Group of the State Council for the Seventh National Population Census, 2020). It can be seen that these provinces have a high proportion of ethnic minority populations apart from the Han, the largest population in China.

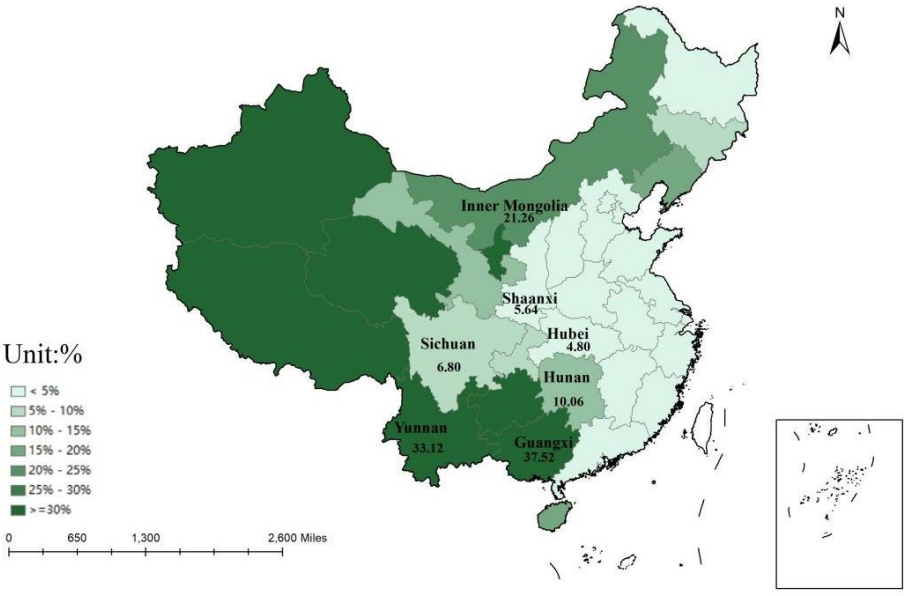


Figure 5. Proportion of ethnic minority population in China (2020)

Due to the small proportion of these ethnic groups, the emergency information is primarily published in Mandarin, the official Chinese language. For some ethnic minority residents, they may not speak or understand Mandarin clearly. This is a great challenge in the information communication on evacuation when a wildfire occurs. The diversity of ethnic minorities in these regions presents a requirement for the accurate information in their native languages.

Challenges for Communication of Evacuation information in Multi-ethnic Communities

According to the Shannon-Weaver model, the process of information communication can be divided into three

main stages: the generation and release of information, the transmission of information and the perception of information. When a wildfire occurs, there are some challenges for government departments to release evacuation information to residents in multi-ethnic communities. According to the process of information communication, the challenges are sorted as follows.

Generation of Information

At this stage, the government needs to generate a reasonable evacuation strategy for the public based on the real - time fire situation. In multi-ethnic areas, there is also a need to express in multiple languages. The challenges at this stage are:

- Dynamic variability of evacuation information. The accuracy of evacuation information is closely related to the real-time state of the fire. The development of the wildfire is dynamic, so it is necessary to dynamically update the evacuation information and release it to the public according to the real-time state, rather than a single information. For example, as the fire grows, the evacuation routes, available shelters, and available safe evacuation time in the information need to be modified.
- Regional difference of evacuation information. In the event of a wildfire, the risk exposure of residents in different regions is different, and in order to avoid possible evacuation congestion, the evacuation information (especially the evacuation routes) in different areas should be different.
- Cross-language expression of evacuation information. In multi-ethnic communities, considering the native language of different ethnic groups, it is also necessary to pay attention to the multilingual expression of evacuation information to meet the information requirements of different ethnic groups.

Transmission of Information

After the evacuation information is generated, the government will release the information through TV, radio, SMS, online social media and so on. Generally, when a wildfire occurs, government authorities use three main channels to release evacuation information to the public in the digital era: (a) Cable broadcasting and television (these two carriers are often reported together as an indicator in China). The emergency information was released through news bulletin or scrolling marquee in the television. In this channel, the length of the information is limited. (b) SMS (Short Messaging Service). When emergencies occur, the government departments compile the emergency information, and the mobile phone operator will release it to users through SMS (Short Messaging Service). This channel also has a limit on the number of words, but it supports the partitioned release of information, that is, the information for different regions can be targeted. (c) Online social media. Emergency information is released on online social networks by official accounts. The advantage of this channel is that there are fewer restrictions on the length of the information.

The popularity of these information carriers is investigated, and their penetration in the multi-ethnic areas with high risk of wildfire is shown in Figure 6, according to the statistics (National Bureau of Statistics, 2022; China Internet Network Information Center,2022). The penetration condition can represent the emergency information coverage in emergency scenarios to some extent.

It can be observed that the mobile phone and Internet are becoming mainstreams, while the penetration of traditional channel, the cable broadcasting and television are relatively low. The popularity of mobile phones is relatively high, and the rates in these high-risk provinces are all over 100%, which means that the use of SMS can achieve full information coverage statistically. However, it should be noted that the mobile phone penetration rate is calculated as the average number of mobile phones per 100 people, and some citizens have more than one mobile phone, while some may not have a mobile phone (especially for the elderly). So, it is improper to release information only with SMS. As for the Internet, the penetration rate in each province is more than half, but there are also differences among provinces, which is related to the level of economic development of the region. Regions with a low level of economic development, such as Yunnan Province, have low Internet penetration relatively.

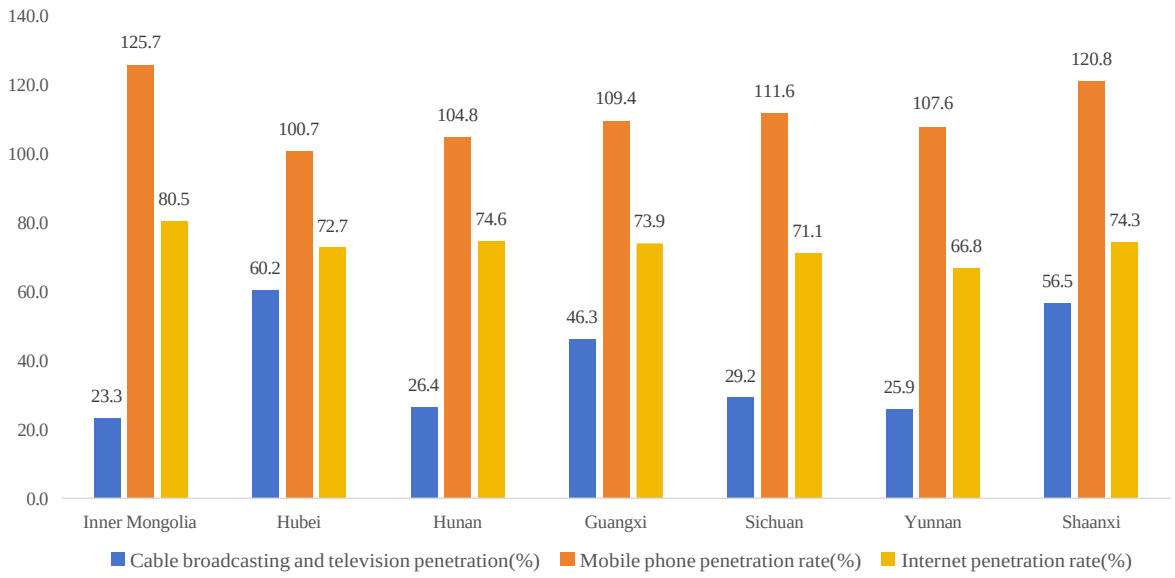


Figure 6. Penetration of Information Carriers in several provinces of China

To summarize, the challenges at the stage of information transmission are:

- Inadequate communication capability of evacuation information carriers. Unlike developed countries, the level of informatization in many developing countries, including China, still needs to be improved. The above analysis of penetration of several information carriers shows that the coverage and spread speed of information carriers in China are still insufficient. In the event of a wildfire, it is likely that not all the residents can receive evacuation information in time.
- Instability of communication infrastructures. In the case of wildfires, it is necessary to take into account the damage to communication infrastructures caused by the fire. Many communication infrastructures, such as the cell phone towers, are located in the wild, and they may be burned when a wildfire occurs, leading to the inability to receive the evacuation information from SMS in the local area.

Perception of Information

After receiving evacuation information through various channels, the residents still need to accurately understand the evacuation information. Apart from the individual differences, the following challenges still exist for the perception of evacuation information in multi-ethnic communities:

- Comprehension impairment due to non-native language or formalization. The evacuation information is primarily published in Mandarin, so the ethnic minority residents may not understand it clearly and rapidly if Mandarin is not their native language. Except for the language barrier, the difference between official expression and the colloquial expression will also make the public understand the emergency information differently. For example, the official naming of specific roads and locations mentioned in the evacuation information may be different from the naming of local residents, which will lead to a sense of unfamiliarity and misunderstanding of evacuation information for the public.
- Irrational perception in crisis situations. When a wildfire occurs, the evacuation information may be selectively ignored and filtered by the residents. Neuroscientists believe that the mechanism by which the human brain selectively ignores and filters information is closely related to the state of activity of neuronal networks. When the brain perceives a threat or an emergency situation, the state of activity of the neuronal network changes in order to better process and transmit information related to these situations. In a wildfire emergency, when the public receives the evacuation information, they may realize that they are in danger and "need to evacuate", but "how to evacuate" (the information on evacuation routes) mentioned in the evacuation message may be ignored, and the evacuation routes may be chosen according to personal experience.

MODEL AND SUGGESTIONS

On the basis of Shannon-Weaver Model, considering the characteristics of wildfire development and information

communication in multi-ethnic communities, a whole-process communication model of evacuation information in multi-ethnic communities is constructed as Figure 7, which also summarizes the challenges and possible responses in each segment. In this model, the official emergency evacuation information publisher, namely the local government or fire department, is considered to be the sender, which generate and send out the evacuation information to be transmitted; and the receivers are the residents in the wildfire-affected regions, who will receive the evacuation information during a wildfire. And the intermediate process from encoder to decoder can be regarded as the process of information transmission, including the channels of television, radio, SMS, online social media and so on. In addition, noise will occur in the process of information transmission, mainly in the form of unofficial and untrue rumors, as well as the interference to the accuracy of information by the dynamic wildfire condition.

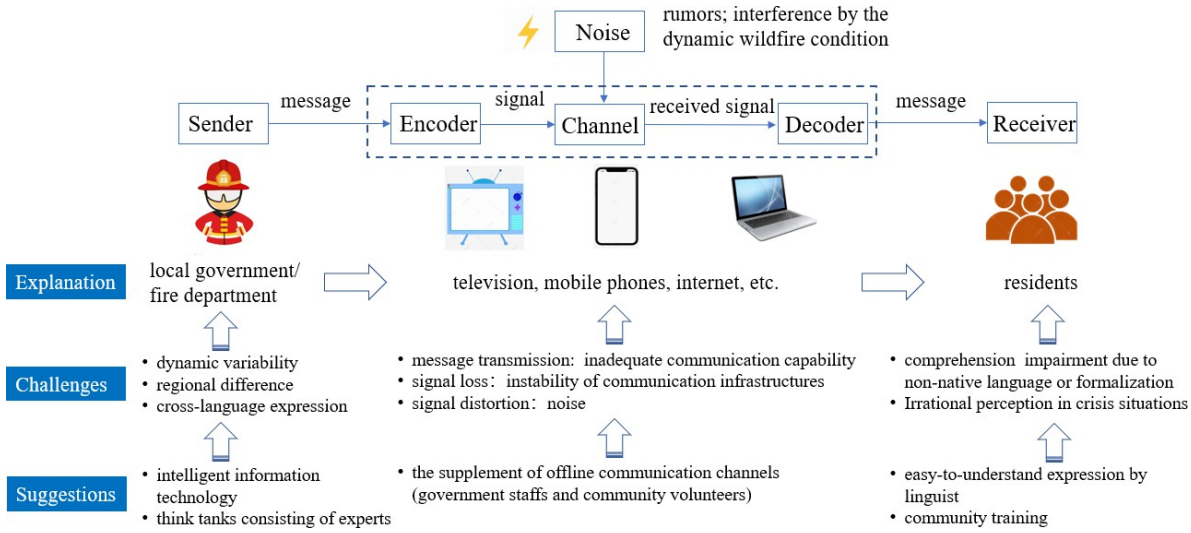


Figure 7. Whole-process communication model of evacuation information in multi-ethnic communities

In the process of communication, various challenges can affect specific segments, as evidenced by the following: when generating evacuation information from the sender, the dynamic variability and regional difference of evacuation information need to be considered, as well as the cross-language expression; in the process of information transmission, the communication capability varies among different information carriers (television, mobile phones, internet, etc.), and there may be regional signal loss or interruption due to damage of communication infrastructures, as well as information distortion due to the addition of noise; when the public receives the information, barriers to information comprehension due to non-native language, formalization, and the problem of irrational perception in crisis situations may make it hard to accurately and fully understand the evacuation information.

In response to these challenges, the authors make some suggestions:

First, the application of intelligent information technology. At present, with the development of artificial intelligence (AI) technology and the optimization of Large Language Models (LLM), the accuracy and speed of machine translation are constantly improving, and these intelligent technologies can be used to assist the release of information in emergency situations. Intelligent machine translation can utilize machine learning, deep learning, and other technologies to train models to recognize multilingual information and automatically generate the appropriate content for information release. With the help of intelligent machine translation, the time for translation of information can be reduced, ensuring the timely and rapid release of information, and the multi - language information can improve the comprehension of multi-ethnic receivers.

Second, the improvements to the crisis response system. China's crisis response system focuses on government-led and multi-party participation, relying on modern technological means and emphasizing early warning and prevention, rescue and relief, and post-disaster recovery, in order to safeguard the safety of public life and property and social stability. However, there are still some deficiencies in the way the current crisis response system operates for the issue of wildfire. For example, there is a lack of timely updated evacuation information and a lack of scientific basis for decision-making, all of which will affect the effectiveness of wildfire response in China. Therefore, we propose to start improving the way the system operates from the following two perspectives in order to increase its efficiency: a) strengthening the evacuation training of the community. Frequent training could help to reduce irrational behaviors of residents in critical situations, so that they can understand the evacuation information more calmly and accurately. b) Establishing organization of think tanks consisting of experts on fire

science, psychology and linguistics. The experts will cooperate with government departments and support decision-making in emergency scenarios, so that the dynamic scenarios of wildfires can be judged more accurately, and the evacuation information can be released and updated with a reasonable frequency and easy-to-understand expression.

Third, the supplement of offline communication channels. Many elderly people seldom use digital communication tools (mobile phones, laptops, etc.), especially for those from ethnic minorities, causing the barriers to receive emergency information. In addition, those with audio-visual impairments are unable to receive information in ordinary form. For these groups, offline face-to-face communication by government staffs and community volunteers is crucial. Meanwhile, the offline channels convey more accurate information, as the information provided by the government staffs and volunteers is official, avoiding the interference of possible rumors (noise). Therefore, despite the obvious disadvantage of inefficiency, the offline channels of information communication remain irreplaceable. In particular, in extreme cases where communication infrastructures are completely destroyed and wide-scale communication cannot be achieved through communication tools (mobile phones, telephones), door-to-door service by the government staffs and volunteers in the communities will be the most reliable form of information communication. At the same time, by carrying shouting or broadcasting equipment, drones can convey instructions, safety alerts and other information to the affected people, improving the public's ability to respond to emergencies.

CONCLUSION

The timely and accurate evacuation information is crucial for the residents during a wildfire. In China, the regions with high wildfire risk are also inhabited by multi-ethnic people, posing some challenges to the communication of evacuation information. In this research, the specific challenges are analyzed with a whole-process communication model, from the generation to the perception of evacuation information.

As a supplement to other research on the public evacuation behavior and decision, the challenges during the information communication are pointed out in this work, and some practical suggestions are also provided. In this work, the study areas are the multi-ethnic areas with high risk of wildfire in China, but the research is also applicable to other areas around the world, as some areas of the world with high wildfire risk are also inhabited by multi-ethnic residents. So, the discussion is helpful and can be served as a reference for the release of information in wildfire situations. Further research on how to solve these problems are still needed, especially on how to provide evacuation with regional difference and cross-language expression, and quantitative analysis on comprehension impairment due to non-native language and public perception in crisis situations.

Due to time constraints, it is not possible to obtain sufficient data at this time. However, we are conducting field studies to collect more specific data. We also plan to conduct an in-depth analysis using more quantitative communication science models, which will help to more accurately assess and explain the issues involved in the study.

ACKNOWLEDGMENTS

This work was supported by GuangDong Basic and Applied Basic Research Foundation [grant number 2023A1515110301&2022A1515110347]; Guangdong Philosophy and Social Science Planning Project [grant number GD23YGL36]; and the Fundamental Research Funds for the Central Universities [grant number 23JNQN09].

REFERENCES

Beloglazov, A., Almashor, M., Abebe, E., Richter, J., & Steer, K. C. B. (2016). Simulation of wildfire evacuation with dynamic factors and model composition. *Simulation Modelling Practice and Theory, 60*, 144-159.

Beverly, J. L., & Bothwell, P. (2011). Wildfire evacuations in Canada 1980–2007. *Natural Hazards, 59*, 571-596.

Cao, Y., Boruff, B. J., & McNeill, I. M. (2016). Is a picture worth a thousand words? Evaluating the effectiveness of maps for delivering wildfire warning information. *International Journal of Disaster Risk Reduction, 19*, 179-196.

China Internet Network Information Center. (2022). *The 52nd Statistical Report on China's Internet Development.* https://www.cnnic.net.cn/NMediaFile/2023/0908/MAIN1694151810549M3LV0UWOAV.pdf

Cohn, P. J., Carroll, M. S., & Kumagai, Y. (2006). Evacuation behavior during wildfires: results of three case studies. *Western Journal of Applied Forestry*, *21*(1), 39-48.

Cova, T. J., Theobald, D. M., Norman, J. B., & Siebeneck, L. K. (2013). Mapping wildfire evacuation vulnerability in the western US: the limits of infrastructure. *GeoJournal*, *78*, 273-285.

Li, L., Ma, Z., & Cao, T. (2021). Data-driven investigations of using social media to aid evacuations amid Western United States wildfire season. *Fire Safety Journal*, *126*, 103480.

McCaffrey, S. M., Velez, A. L. K., & Briefel, J. A. (2013). Differences in information needs for wildfire evacuees and non-evacuees. *International Journal of Mass Emergencies & Disasters*, *31*(1), 4-24.

McCaffrey, S., Wilson, R., & Konar, A. (2018). Should I stay or should I go now? Or should I wait and see? Influences on wildfire evacuation decisions. *Risk analysis*, *38*(7), 1390-1404.

McLennan, J., Ryan, B., Bearman, C., & Toh, K. (2019). Should we leave now? Behavioral factors in evacuation under wildfire threat. *Fire technology*, *55*, 487-516.

National Bureau of Statistics. (2022). *China Statistical Yearbook*. China Statistics Press, Beijing.

National Forestry and Grassland Administration. (2012-2021). *China Forestry Statistical Yearbook*. China Forestry Press, Beijing.

Office of the Leading Group of the State Council for the Seventh National Population Census. (2020). *China Population Census Yearbook*. China Statistics Press, Beijing.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.

Stasiewicz, A. M., & Paveglio, T. B. (2021). Preparing for wildfire evacuation and alternatives: Exploring influences on residents' intended evacuation behaviors and mitigations. *International Journal of Disaster Risk Reduction*, *58*, 102177.

Steelman, T. A., McCaffrey, S. M., Velez, A. L. K., & Briefel, J. A. (2015). What information do people use, trust, and find useful during a disaster? Evidence from five large wildfires. *Natural Hazards*, 76, 615-634.

Stidham, M., Toman, E., McCaffrey, S., & Shindler, B. (2011). Improving an inherently stressful situation: the role of communication during wildfire evacuations. In *Proceedings of the second conference on the human dimensions of wildland fire* (pp. 96-103). Gen. Tech. Rep. NRS-P-84.

Toledo, T., Marom, I., Grimberg, E., & Bekhor, S. (2018). Analysis of evacuation behavior in a wildfire event. *International journal of disaster risk reduction*, *31*, 1366-1373.

Vaiciulyte, S., Hulse, L. M., Veeraswamy, A., & Galea, E. R. (2021). Cross-cultural comparison of behavioural itinerary actions and times in wildfire evacuations. *Safety science*, 135, 105122.

Vélez, A. L. K., Díaz, J. M., & Wall, T. U. (2017). Public information seeking, place-based risk messaging and wildfire preparedness in southern California. *International journal of wildland fire*, 26(6), 469-477.

Wang,Y.C., Ma, J.X., Liu, Y.H., & Bai, Y.J. (2021). The evacuation model of forest tunnel fire based on the characteristics of smoke diffusion. *Journal of Nanjing Forestry University (Natural Sciences Edition)*, *45*(6): 177−184.



IT Solutions for Crisis Management

Track Chairs

Jürgen Moßgraber, Fraunhofer IOSB

Anastasios Karakostas, DRAXIS Environmental

Bas Lijnse, Netherlands Defense Academy

Drone Swarm for Post-Wildfire Hot Spot Detection: Technology Assessment and PoC Demonstrator

Micha Sende*
Lakeside Labs GmbH, Klagenfurt, Austria
sende@lakeside-labs.com

Christian Raffelsberger
Lakeside Labs GmbH, Klagenfurt, Austria
raffelsberger@lakeside-labs.com

Samira Hayat
Lakeside Labs GmbH, Klagenfurt, Austria
hayat@lakeside-labs.com

Armin Köfler
JOANNEUM RESEARCH
Forschungsgesellschaft mbH, Digital,
Graz, Austria
armin.koefler@joanneum.at

Alexander Almer
JOANNEUM RESEARCH
Forschungsgesellschaft mbH, Digital,
Graz, Austria
alexander.almer@joanneum.at

ABSTRACT

Post-wildfire hot spots are a severe problem for fire fighters since they are hard to detect and can reignite fires. The tedious search can be greatly improved by using drones. However, developing and controlling an autonomous drone swarm is a challenging task. An assessment with domain experts reveals the suitability of drone swarms for hot spot detection. We develop a heterogeneous multi-drone system that can support fire fighters by creating thermal and visual aerial images and automating the hot spot detection process. During a joint exercise with a fire fighter brigade we demonstrate that the proposed system is able to autonomously search, detect, and report hot spots. Therefore it is able to speed up the search process while reducing the required number of personnel.

Keywords

Drone Swarm, UAV Swarm, Wildfire Hot spots, Technology assessment, PoC Demonstrator

INTRODUCTION

The critical nature of disaster response missions make them ideal use cases for drone swarms. With climate change, the number and intensity of natural disasters is rapidly increasing. According to the "World Disasters Report – Trends in Disasters" 2022, the number of geological disasters increased from 19 in 2020 to 33 in 2021 (International Federation of Red Cross and Red Crescent Societies, 2022). According to a European Commission's report on wildfires, the number of wildfires in the EU increased by 48 % from 2021 to 2022, while the affected area increased by 86 % (Joint Research Centre, 2023). With an increase in number and intensity, wildfires are causing an increased havoc on the environment, human health, and infrastructure (United Nations Environment Programme, 2022).

Fighting wildfires requires a coordinated effort by fire fighters, emergency management, and specialized aircraft teams. In recent years, drones also known as unmanned aerial vehicles (UAVs) equipped with visual and thermal

*corresponding author

sensors have been used to support these efforts. Drones offer a cost-effective solution, compared to traditional methods for visual overview, such as using satellite imagery or aerial photography using helicopters. Fire departments equipped with drones may have a lot more flexibility in terms of operation time to generate and analyze such overview images. Additionally, drone swarms reduce the overall mission time significantly when compared to search using watch towers, ground robots, or ground personnel.

After wildfires are extinguished to a significant degree, fire fighters must search the affected area for hot spots that may be hidden inside burned tree trunks or buried out-of-sight. Sparks flying from these hot spots may lead to future fires. Therefore, the post-disaster phase of detecting hot spots, which may span over multiple weeks, is very important to prevent reignition. Currently, hot spot detection is mostly performed manually by fire fighters that are equipped with handheld thermal cameras. They may be supported by drones which are manually steered and do not provide automatic hot spot detection. These efforts require a lot of personnel. In this paper we describe a multi-drone system to aid fire fighters in hot spot detection in the post-disaster phase of wildfires.

Our proof of concept (PoC) system implementation uses a fixed-wing drone equipped with a visual camera, four multirotor drones equipped with thermal cameras, visual cameras, and communication interfaces (WiFi mesh and LTE/5G), two ground control stations (GCSs) for human interaction, and three servers for computing and data storage. The system is scalable in terms of the number of drones, and is designed with clear interfaces between humans and drones, while requiring minimum intervention from humans.

The system design takes into account feedback from first responders and emergency management experts gathered through questionnaires and expert interviews. It is noted from the feedback that current usage of drone imagery for hot spot detection puts a strain on the resources, requiring two persons per drone, a pilot and one person to analyze the images. The developed PoC by design requires two persons for the operation of the entire swarm. The purpose of this work is to demonstrate that an autonomous swarm is a feasible solution in terms of optimal resource usage in post-wildfire hot spot detection. The system is able to autonomously identify hot spots from the air over large areas, leaving the fire fighters free to tackle the task of extinguishing the fires. The focus of this paper, however, is not to provide quantitative results comparing the efficiency of manually-operated versus autonomous drones for hot spot detection.

The mission starts with the fixed-wing drone flying autonomously to generate a real-time aerial overview image of large areas. This provides the ground personnel with an up-to-date situation overview image to identify regions of interest (ROIs) for closer inspection. The ROIs can then be assigned to one or more multirotor drone swarms, which cover less area but provide greater detail. A single pilot can track the swarm status using a GCS. To synchronize takeoff in a swarm, the swarm pilot gives the mission start command. The swarm takes off, flies, searches, and lands autonomously. During search, the drones detect hot spots and transfer images and hot spot information to the ground, where they are geo-referenced. The location and intensity information of the hot spots is passed to the fire fighters who take necessary action to extinguish them.

Communication plays a critical role for successful swarm operation during the mission. The drones must avoid collisions autonomously. The swarm pilot must be able to track the swarm mobility and the hot spot images, and information must arrive in real-time manner at the GCSs. To support actions that benefit from device-to-device connectivity, e.g., collision avoidance, we use WiFi mesh; to support high-throughput real-time image transfer, LTE/5G networks are used. Thus, we are able to demonstrate a fire fighting mission aided by a drone swarm which makes the best use of two different communication technologies.

Our main contribution is the demonstration of a distributed swarm robotic system integrated from several sub-systems in a realistic environment including stakeholders. Compared to state of the art drone operations in wildfire scenarios, our system reduces the required ground personnel count while increasing the speed at which the situation overview is generated.

The rest of the paper is organized as follows: First, we list related work and report the outcome of the technology assessment. Then we describe the system architecture and its components, followed by details of the proof of concept demonstration and a conclusion.

RELATED WORK

There is no shortage of research work focusing on drone swarm or multi-drone use for search and rescue (SAR) (Arnold et al., 2020; Dominguez et al., 2017; Dong et al., 2021; Guo et al., 2019; Mbaitiga & Shosaku, 2022; Quan et al., 2019; Scherer et al., 2015; Steinha¨usler & Georgiou, 2022). A small number of these focus on real-world system implementation (Dong et al., 2021; Lauterbach et al., 2019; Quan et al., 2019; Scherer et al., 2015; Steinha¨usler & Georgiou, 2022; Terzi et al., 2019). Some work on system implementation highlights the

necessity for seamless integration of different technologies (Scherer et al., 2015; Tinetti & Riveros, 2021), which is a challenging task. The task becomes further complicated when using an autonomous drone swarm. In this paper, we present a multi-drone system to aid fire fighters. The drones fly completely autonomous to search an area and detect hot spots while avoiding collisions.

A drone swarm may support a team of fire fighters in many ways. One of the biggest advantages of using drone swarms is their ability to generate large-scale, near real-time, visual situational awareness by 3D area mapping (Lauterbach et al., 2019). The drones may additionally collect other valuable sensor data such as the levels of hazardous plumes in the air and reduce the risk to the fire fighters' lives (Seiber et al., 2018). Drone swarms are also ideal for victim detection and localization (Dong et al., 2021; Quan et al., 2019; Scherer et al., 2015; Steinha¨usler & Georgiou, 2022) as they provide a much faster area search as compared to ground personnel or robots. Finally, they help locate sources of disaster such as fires or toxic plumes (Seiber et al., 2018; Viseras et al., 2019). In our proposed system, the autonomous drone swarm locates hot spots in the post-disaster phase. Locating and extinguishing the hot spots mitigates the risk of future fires.

The increasing number of wildfires in recent years has seen an increased interest among the drone research community to develop systems aiding fire fighters. Pasini et al. developed a framework for cooperation between aerial and ground robots (Pasini et al., 2022). A drone is used to gather a visual overview of the area of interest while identifying hot spots. The overview image is transferred to a ground station, which generates a path to the identified hot spots for a rover. Viseras et al. performed a measurement campaign with three drones to collect visual and thermal images of wildfires and hot spots (Viseras et al., 2019). Hot spots with a minimum size of 15 cm could be identified during image analysis from a height of 90 m. Unlike the above two works, in our system, the drones are equipped with visual and thermal cameras to autonomously identify hot spots in a real-time manner during flight. Dong et al. used a similar approach for survivor detection post-disaster using YOLOV3-MobileNet on a single drone (Dong et al., 2021). However, the drone swarm system implementation was not the focus.

The agents in a drone swarm may make autonomous flight decisions either using visual information (Schilling et al., 2022) or using communication (Scherer et al., 2015). In this work, we rely on WiFi mesh communication for flight decisions due to the small drones' payload limitations. The drones are also equipped with 5G modems for sensor data transfer. The use of 802.11s mesh network to operate a drone swarms has been proposed in (Scherer et al., 2015), with related measurement campaigns performed in (Hayat et al., 2015). It was shown that in a multi-sender scenario (two drones generating traffic and sending to a ground station), the network offered fairness and an average throughput of 10 Mbit/s up to a distance of 450 m from the ground station. Scalability of wireless mesh networks is a known concern, mostly caused by link sharing (Lei et al., 2022). How mesh technologies such as WiFi fare in real-world drone swarms with more than ten agents depends on the amount of traffic needed to be exchanged among the agents, and is still an open topic.

For the transfer of high-throughput real-time payload data such as visual and thermal images, 5G networks are ideal. Experimental campaigns report drone-to-ground station throughput of up to 50 Mbit/s, with latency as low as 47 ms in a suburban environment (Fakhreddine et al., 2022). The average latency has been shown to vary with signal-to-interference and noise ratio, with its variance strongly impacted by handover frequency in both urban and suburban scenarios in (Luo et al., 2022).

TECHNOLOGY ASSESSMENT

In order to identify the most important application scenarios for a drone swarm and to gather initial requirements, we asked first responders and emergency management experts from Austrian ministries to complete a questionnaire. In total 33 experts replied. The majority (57.6 %) have more than 15 years of experience, 42.4 % are in senior command, 30.3 % in mid-level command and 24.2 % in operational service. Such a technology assessment is important, since it helps to determine the feasibility and suitability of the system for the specific needs and requirements of the first responders. The questionnaire introduced the general idea of autonomous drone swarms, equipped with different sensors, to support the emergency response. It contained questions to assess the experts' opinions concerning the suitability of such a system in different scenarios and specific use cases, the anticipated risks, dangers, advantages and disadvantages, both with closed Likert scale and open answers.

Due to space constraints, we only provide selected results of the user survey. In a series of questions, we asked the experts whether they a) believe that such a system introduces risks, b) disturbs the work of first responders, and c) brings them more disadvantages than advantages. Figure 1 shows the results for four application scenarios: wildfires, other natural disasters, person monitoring (e.g., at large events), and monitoring critical infrastructure. The majority of the experts agreed that a drone swarm introduces risks in person monitoring and wildfire scenarios. Most experts disagreed that such a system introduces risks for other natural disaster responses and monitoring

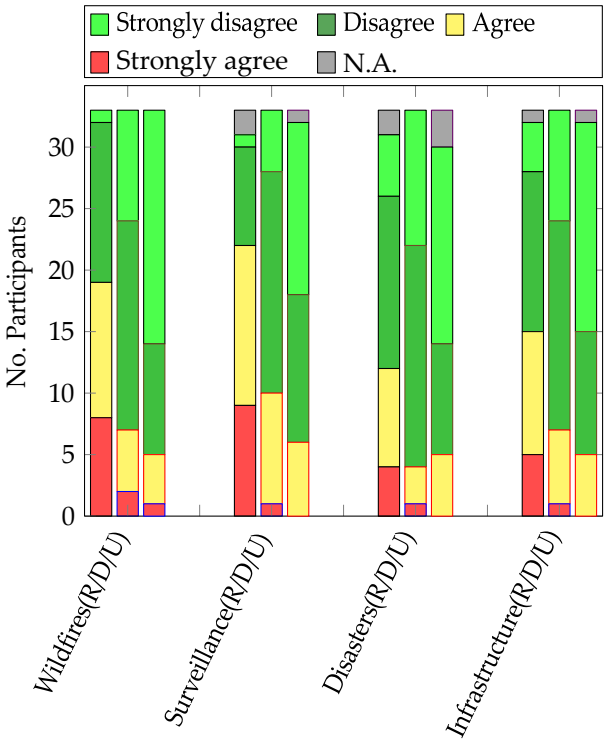


Figure 1. Does a drone system introduce more risks (R), dangers (D), or more disadvantages than advantages (U) to the use cases?

of critical infrastructure. The majority disagreed that a drone swarm disturbs the operation or introduces more disadvantages than advantages. It is safe to say that drone swarms are seen as beneficial in all scenarios.

We also asked the experts about specific use cases for drone swarms in the wildfire scenario, such as: detection of wildfires (e.g., after lightning strikes), monitoring the actual wildfire response, hot spot detection, and the post-disaster damage assessment. Figure 2 shows that hot spot detection and damage assessment are seen as the most suitable use cases to be supported by drone swarms.

The experts were also asked to identify the three most important/useful and the three most problematic/risky/negative factors when using such a system. Many experts mentioned cost as an important factor, both as a strength and as a weakness. Some experts were concerned with the initial investment costs, while others saw the potential for future cost savings due to a reduction in the number of first responders required. Similarly, the complexity of the system was seen as a crucial factor. While some experts were concerned with initial and ongoing training needs to operate the system, others saw a high degree of autonomy as a potential. Legal restrictions on the operation

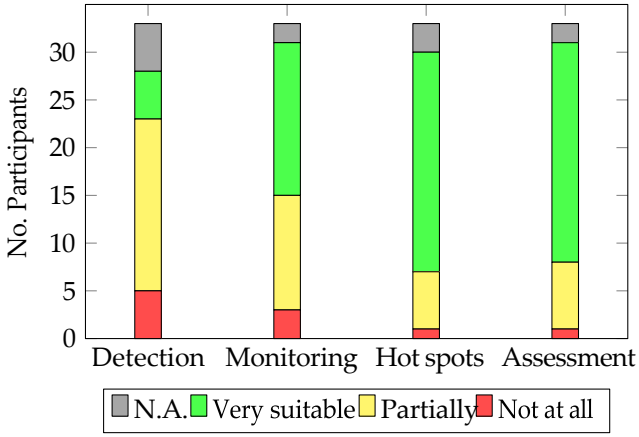


Figure 2. Is a drone swarm suitable for specific use cases in a wildfire scenario?

of autonomous drones and the impact of weather were seen as the biggest risks. Improved assessment times and situational awareness were cited as the biggest potential benefits.

Subsequent to the PoC demonstration, an evaluation was conducted to collect feedback from the first responders. The outcomes of this evaluation are elaborated in the proof of concept section.

Requirements

The experts were also asked about requirements towards such a system. The answers included: ease of use, both when controlling the system and when analyzing the gathered data; providing data interfaces to existing systems; well-defined safety regulations to ensure the safety of the personnel on ground and air forces. Furthermore, they mentioned important aspects such as flexibility (e.g., adapt the size of the swarm), quick setup, real-time monitoring, and robustness. Interestingly, there were conflicting opinions about the costs of such a system. While some experts see such a system as a way to reduce the response costs, other expect the system to increase costs.

SWOT Analysis

Based on the answers we performed a SWOT analysis with the following results.

Strengths: The system is seen to handle resources efficiently and hence to reduce cost regarding acquisition, operation, and maintenance. It can increase the speed of a mission, create increased situational awareness, and allow early detection of incidents. Its versatility allows it to be adapted and used for different applications.

Weaknesses: The complexity of the system is seen as its main weakness. This complicates the integration into existing systems, hence limiting its acceptance. Due to the complexity it might also be prone to failures and require regular maintenance, e.g., in changing weather conditions or crowded air space. Operating such a complex system requires training which could drive up costs, while current legislation does not allow a single swarm pilot but requires one pilot per drone.

Opportunities: The system can exploit already existing interfaces of other systems as well as existing, well-defined safety regulations to boost its acceptance. It can play out its flexibility in the challenging and changing conditions of rescue operations.

Threats: Outside influences that can threaten the system operation include incompatibilities with other systems, badly educated personnel, hacking attacks, external disturbance (e.g. interference on the communication channels), and mid-air collisions.

Concluding the assessment, we can say that most experts agree that drone swarms are a suitable technology for hot spot detection in wildfire scenarios. Therefore, we target this scenario with our drone swarm system. To address the concerns regarding system complexity we design our system with a focus on autonomy. Most steps, including mission execution, swarm coordination, communication, image analysis, and feature detection are automated to ease the system operation. This includes a collision avoidance algorithm to minimize the aforementioned risk of collisions in the air.

SYSTEM

In this work we propose a system consisting of several sub-systems that are interconnected using different technologies. An overview is given in Figure 3.

System Components

The proposed system can be organized into three technical sub-systems plus a user interface (UI) for interaction with users, i.e., pilots and fire fighters. The sub-systems are: the fixed-wing drone, the multirotor drone swarm and the server infrastructure for off-board analysis and storage.

The first subsystem consists of a fixed-wing drone (Skywalker EVE-2000) equipped with a visual camera (Sony IMX249 sensor with 2.3 megapixel mounted on a 1D gimbal for roll compensation) that creates an aerial image, a map server based on the open source implementation GeoServer[1] to store that image, and a custom build GCS that controls the mission of the drone and visualizes the progress (Fanta-Jende et al., 2023). The drone is controlled by a Pixhawk 2 flight controller running the Ardupilot firmware and localizes using a Cube Pilot Here3 RTK GPS. The onboard computation is performed by an NVIDIA Jetson Xavier NX. The drone has a maximum takeoff weight of 4.6 kg and a cruise speed of 75 km/h.

[1] https://geoserver.org/

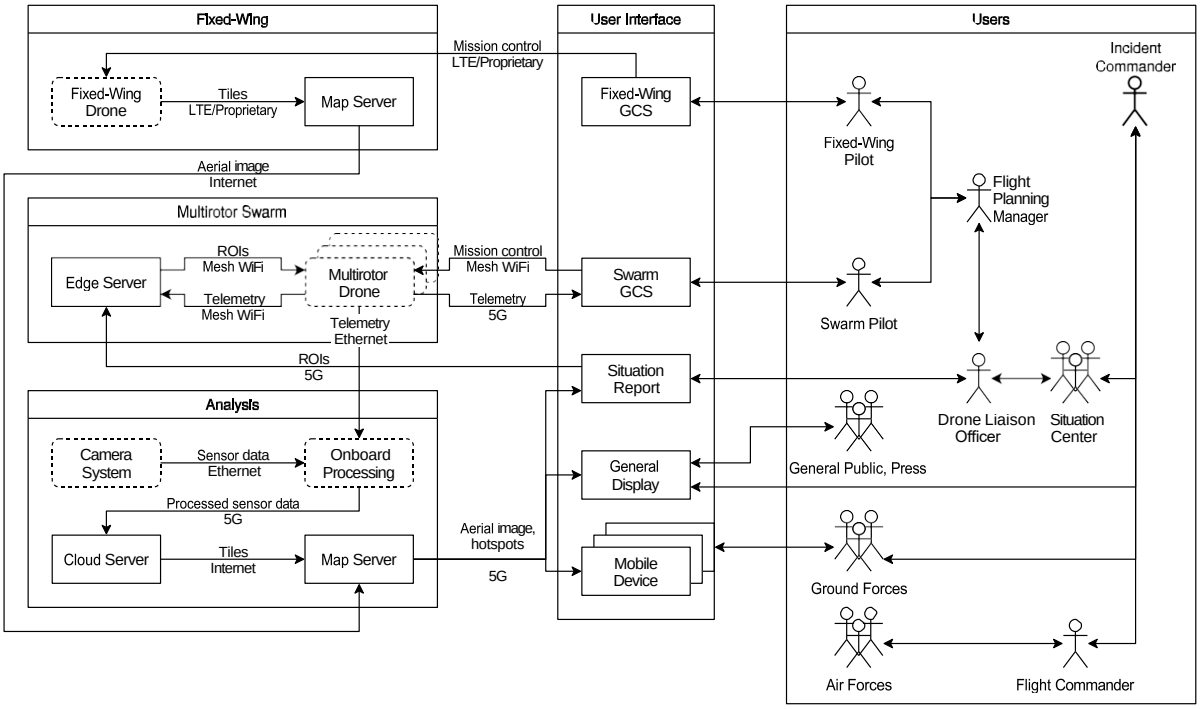


Figure 3. System Architecture with aerial components marked by dashed lines.

The second sub-system, the multirotor drone swarm, comprises several multirotor drones (twinFOLD KAT, twinFOLD GEO, twinFOLD SCIENCE)[2], an edge server that computationally supports the drone swarm, and a GCS for controlling the drone swarm. The drones are controlled by algorithms running in the Robot Operating System (ROS) on a Raspberry Pi 4 B connected to a Pixhawk flight controller that runs the Ardupilot firmware.

The third sub-system includes the components related to image analysis. Its main task is to locate the hot spots and provide a map server that manages the recorded images and the results of the hot spot detection. Onboard the multirotor drones are the following sensors: an RGB camera (Ximea CMV20000), two thermal infrared cameras (custom built by LEADER Photonics) and a GNSS/INS sensor (SBG Quanta). An NVIDIA JETSON AGV Xavier board is used for controlling the sensors and pre-processing sensor data onboard the drones. A custom built trigger box ensures that all sensor data (images and pose) is collected synchronously. The setup is shown in Figure 4. The takeoff weight including batteries for a flight time of roughly 15 min is around 14 kg. It should be noted, that the component weights have not been optimized, hence providing significant potential to extend the flight time. For further analyzing the sensor data, a cloud server allows computing intensive post-processing tasks.

The user interaction happens through the UI sub-system. Both fixed-wing drone and multirotor swarm have a GCS to control and monitor the mission. The main mission results are presented through the situation report terminal. It is a web front end that shows base maps, the real-time aerial image collected by the fixed-wing drone, the locations and state of hot spots and the location of fire fighters' ground forces and multirotor drones. It can also be used to define and send ROIs to be mapped by the swarm or to send commands to first responders that carry smart phones with a dedicated Android app. Additionally, there can be a non-interactive terminal (general display) to inform the general public about the mission.

The user side reflects the hierarchy of the fire fighter personnel. The main interaction points to the technical system are the pilots (fixed-wing and swarm) which are coordinated by the flight planning manager. He interacts with the drone liason officer who gets his information from the situation report terminal and gives instructions for the drone missions.

Communication

The system components are interconnected using different technologies. Communication between the fixed-wing drone and the GCS may use different technologies (Fanta-Jende et al., 2023). In most situations, 5G/LTE mobile networks provide sufficient bandwidth and a stable connection for transmitting the payload data to the GCS. For

2https://www.twins.co.at/en/multirotorsystems/

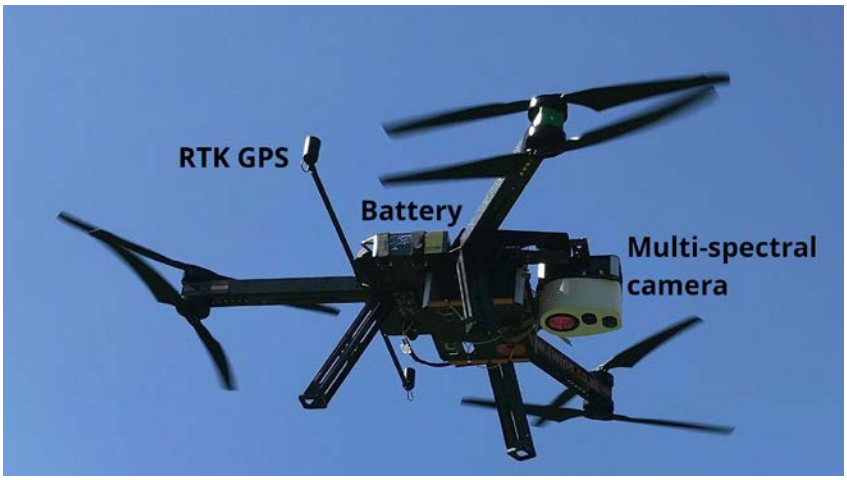


Figure 4. Multirotor drone setup (the processing boards are underneath the drone).

this purpose the drone and GCS are equipped with LTE routers (Teltonika RUT240 on the drone and RUTX11 on the GCS). Alternatively, a WiFi 802.11n network using a tracking antenna or a proprietary data link may be used. This alternative is important, as there may be a lack of mobile network coverage in the area of operation. Communication is used to control the mission from the GCS and transmit the recorded images to the map server on the ground. Coordination with the multirotor swarm is not needed at this stage since the fixed-wing drone mission is separated in time from the multirotor swarm mission.

The multirotor swarm uses a combination of WiFi mesh and LTE/5G mobile networks. Communication and coordination within the swarm is based on WiFi mesh to increase robustness in the absence of infrastructure. The mobile network is used to transmit payload data to the cloud server for processing and returning the result to the ground stations, and for receiving the ROIs to be searched.

Algorithms

The drones are programmed and controlled using ROS version 1[3]. Many of the software components are available as open source and links are given in the following.

A typical mission begins by fire fighters defining the mission area. The fixed-wing pilot enters it into the GCS which is a custom implementation connected to a locally running ROS master. When the fixed-wing pilot starts the mission, the drone autonomously surveys the mission area using its onboard camera to create an aerial image. It records images and performs visual simultaneous localization and mapping (SLAM) to localize itself while creating the aerial image. Being connected to the ROS master on the GCS, it transmits individual tiles of this aerial image to a map server located on the ground allowing the GCS to continuously update the visualization of the aerial image in real-time. This includes the steps of densification and stereo reconstruction[4], surface generation, semantic labeling[5], orthorectification[6], tileing[7][8], and object detection[9] (Fanta-Jende et al., 2023). Once the aerial image is complete, the fire fighters can define ROIs that should be covered in greater detail.

In the multirotor swarm, each drone, the edge server, and the swarm GCS run individual ROS instances that use the CPSwarm libraries[10]. They are connected using the swarmio bridge[11] (Sende et al., 2021) that forwards selected topics using the the ZeroMQ Message Transfer Protocol (ZMTP) over a (wireless) network. This increases the robustness of the system in unreliable networks. Hierarchical finite state machines (FSMs)[12] are used to model and implement complex behaviors controlling the drone from the onboard computer (Sende et al., 2021). The ROIs are received by the edge server that automatically assigns them to the available drones following a first-come,

3 https://www.ros.org/
4 Plane sweep library based on CUDA: https://developer.nvidia.com/cuda-downloads
5 Open neural network exchange: https://github.com/onnx/onnx
6 OpenREALM: https://github.com/laxnpander/OpenREALM
7 Tiled map service: https://wiki.osgeo.org/wiki/Tile_Map_Service_Specification
8 MapTiler: https://epsg.io/3857
9 YOLOv5: https://github.com/ultralytics/yolov5
10 https://github.com/cpswarm
11 https://github.com/cpswarm/swarmio/
12 FlexBE: https://github.com/FlexBE/

first-served policy. As soon as a drone receives an ROI it computes a coverage path following a Bustrophedon back and forth pattern to search the given ROI[13]. The spacing between the different legs of this path is computed considering the characteristics of the camera system, i.e., field of view and altitude above ground. The swarm pilot uses the GCS to monitor the state of the drones and control the mission. He issues commands to start or stop the mission. For safety reasons, the drones wait for the mission start command before taking off.

After take off, a drone moves to the assigned ROI, activates the sensors and begins to autonomously search the ROI. Controlled by the onboard processor, the drone jointly captures thermal infrared and visual images and synchronizes them with its current pose.

The thermal images are analyzed on board and hot spots are detected if the temperature difference to the surroundings passes a given threshold. There can be an upper as well as a lower threshold. Hot spot detection using deep-learning neural networks or other AI approaches were dismissed due to the unavailability of suitable training data sets. The results are very small text-based data samples. Enriched with additional information such as temperature, size, and shape, they are sent to the cloud server for post-processing. The cloud server performs the more computationally expensive tasks of referencing the hot spots spatially and grouping neighboring detections using a morphological opening operator that also cleans up the data to smooth the detected polygons. The hot spots are labeled and linked with the meta data. Over the course of a mission, information about a hot spot is received multiple times in different image frames. This time series information is used for a scene analysis to improve and update the detection results.

The visual image data cannot be processed on board due to limited computing hardware and is instead transmitted to the cloud server for processing. For saving bandwidth, it is compressed losslessly on board the drone using the JPEG 2000 standard to guarantee that the transport does not influence the processing result. The cloud server generates a sharpened, orthorectified aerial image, both visual and thermal, containing the hot spots and their state (e.g., detected, assigned, extinguished). The result is stored on the map server which provides it to the user interface for further processing by the fire fighters.

While we also process the visual images on the cloud server for detection of humans and objects with YOLO deep-learning algorithms, it is not further described in this paper because it is out of its scope.

When a drone finishes the search, it returns, marks the ROI as completed and waits to be assigned to another ROI.

PROOF OF CONCEPT

To demonstrate the functionality of the proposed system, we performed a common exercise with the fire fighters in a closed stone quarry with an area of approximately 10 ha. The focus of the demonstration lies on the drone swarm component of the system which provides the main novelty and carries the greatest challenge. Besides showcasing the system capabilities to the main user group, the reason for the joined demonstration exercise was to test its usability during wildfire fighting missions. We assume the mission area to be a sub-part of the overall incident region surveyed by the fixed-wing drone. Compared to the fixed-wing drone, the multirotor drones cover a smaller area in greater detail. This helps firefighters deployed in the field to capture their direct surrounding at high detail without exposing them to the strain and risk involved in traditional, foot-based reconnaissance. While single-drone systems are already used for wildfire fighting, they are very personnel demanding (for each drone the fire fighters currently require one pilot and one person for mission planning and image observation). Our demonstration showed that a single operator can control four drones simultaneously by simply providing ROIs. Once the ROIs have been sent, the drones operate autonomously, executing their tasks without further human intervention. Due to resource limitations only four drones have been utilized for the demonstration. However, it is fair to assume that this number can be scaled up, as we performed mixed-reality experiments with simulated and physical drones to evaluate the coordination algorithm using up to 16 drones.

For this demonstration, the fire fighters created several hot spots of different types (burning wood or charcoal) shown in Figure 5a. For safety reasons the hot spots were deployed on rocky ground and not within forests. If hot spots provided a lower temperature footprint due to tree or bush cover, the detection threshold could be adapted accordingly.

Following the mission approval by the incident commander, the drone liaison officer defined four ROIs within the mission area using the situation report terminal, see Figure 5b. They cover a region of roughly 2 ha and are located at a distance up to 370 m from the command center. The ROIs were automatically transferred to the edge server and assigned to the four multirotor drones shown in Figure 5c. Once commanded from the flight planning manager, the swarm pilot started the mission using the swarm GCS. The drones computed flight paths as shown in Figure 5e to autonomously search the assigned ROIs. The drones then autonomously took off (Figure 5d) to altitudes between

[13] https://github.com/cpswarm/swarm_functions/tree/noetic-devel/coverage_path

30 m and 60 m and started moving to their designated ROI while autonomously avoiding collisions with other drones. Flying at a speed of 3 m/s and having a 30 ° camera field of view, they could cover an area of approximately 72 m^2/s while detecting all different types of hot spots. The low flying speed was selected for safety considerations, such as accommodating communication delays or outages that negatively impact collision avoidance. However, there is potential to increase the speed, particularly during ROI approach. During flight, the status of all drones could be monitored through the situation report terminal (blue symbols in Figures 5b and 5f) as well as the swarm GCS (red arrow in Figure 5e). The detected hot spots were visualized in the situation report terminal (warning triangle in Figure 5f) and reported to the incident commander. He then assigned them to individual fire fighters using a smart phone app shown in Figure 5g. The fire fighters moved to the location, confirmed (Figures 5h and 5i) and extinguished the hot spots (Figure 5j).

Based on the results of the initial survey and discussions with experts (see the technology assessment section), a second questionnaire was developed. The focus was on social science inquiries as well as the potential benefits of future market-ready systems based on the swarm technology.

The experts identified several advantages of the system. Most experts consider the demonstrated system highly useful because it is anticipated to enable faster detection of hot spots, particularly through its autonomous operation. A (partially) autonomous swarm of drones will fundamentally expand the capabilities of the first responders by providing a larger and more valuable set of real-time mission data. This enables more precise planning of operations, more efficient firefighting procedures, and enhanced protection for emergency personnel, third parties, nature and wildlife. Two advantages have been named: 1) it accelerates the firefighting efforts, as a larger area can be searched in less time and 2) it reduces the need for manpower, as a single pilot can control multiple drones at once. In this sense, the system has been seen as an extension (not a replacement) of first responder personnel, which can be assigned with different tasks. The coordinated mission execution and the precision with which hot spots were detected are seen as an impressive demonstration of the system's strengths. Standard systems transmit images and videos, which lack the support for orientation. The presented system directly determines GPS positions of events (i.e., hot spots on the images are geo-referenced) and can lead emergency personnel straight to the point of interest.

Several disadvantages were identified. A significant risk is the potential for damage to third parties or objects due to faulty behavior. The greatest danger remains an uncontrolled crash, which can cause harm to people and materials. The lack of a legal framework for the deployment of a drone swarm is another weakness, mentioned by the first responders. Drones are already a substantial aid to all emergency services in dangerous and time-critical operations. However, the legislation and product maturity are seen as big challenges to be overcome. For efficient operation, comprehensive training of the emergency personnel is necessary. There is also a risk in a lack of acceptance of such systems by the emergency personnel themselves and in terms of public acceptance (e.g. privacy concerns).

CONCLUSION

In this paper we present a heterogeneous multi-drone system to support fire fighters with post-wildfire hot spot detection. According to a survey among domain experts, hot spot detection is a suitable application for such a system. They see the advantages in reduced mission time and increased situational awareness; the main disadvantage in its complexity. The system is still in an experimental phase and not yet ready to be deployed as a product. The system could be demonstrated in a controlled environment, with operation mainly performed by the research team, indicating a technology readiness level (TRL) of 4. It falls short in usability, reliability and security, which are critical aspects for effective oepration in actual fire fighting missions. Despite these limitations, we were able to demonstrate the system's capability to autonomously search for, successfully detect, and report hot spots in given ROIs. This outcome is encouraging and suggests potential for significant improvements in firefighting missions, particularly in addressing the increasingly pressing issue of forest fires.

ACKNOWLEDGMENTS

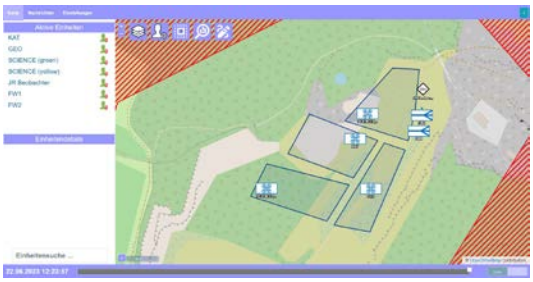
We would like to thank the volunteer fire fighter brigade Gumpoldskirchen, Austria for their support and IFR Ing. Richard Feischl for the task settings and the tactical and operational coordination.

This work received funding by the security research program KIRAS of the Austrian Federal Ministry of Finance (BMF) under grant agreement no. 879682 (UASwarm).

During the development of the communication system and the coordination and control algorithms, we used infrastructure provided by the 5G Playground Carinthia, which is funded by the Carinthian Agency for Investment Promotion and Public Shareholding (BABEG). The 5G Playground is operated by BABEG and financed by means of the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK) and the government of the Carinthian provincial.



(a) Two different hot spots: Charcoal and wood.



(b) ROI definition.



(c) Multirotor drones before takeoff.



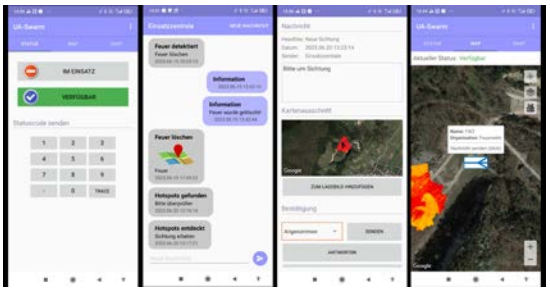
(d) Flying multirotor drones.



(e) Search path of the multirotor drones.



(f) Detected hot spots.



(g) Assigned hot spot on the mobile device.



(h) Hot spot confirmed with thermal camera.



(i) Hot spot confirmed with thermal camera.



(j) Hot spot being extinguished.

Figure 5. Photos and screenshots from demonstration exercise.

REFERENCES

Arnold, R., Jablonski, J., Abruzzo, B., & Mezzacappa, E. (2020). Heterogeneous UAV Multi-Role Swarming Behaviors for Search and Rescue. *Proceedings of the Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, 122–128. https://doi.org/10.1109/CogSIMA49017.2020.9215994

Dominguez, M. H., Nesmachnow, S., & Herna´ndez-Vega, J.-I. (2017). Planning a Drone Fleet using Artificial Intelligence for Search and Rescue Missions. *Proceedings of the International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*. https://doi.org/10.1109/INTERCON.2017.8079646

Dong, J., Ota, K., & Dong, M. (2021). UAV-Based Real-Time Survivor Detection System in Post-Disaster Search and Rescue Operations. *IEEE Journal on Miniaturization for Air and Space Systems*, *2*(4), 209–219. https://doi.org/10.1109/JMASS.2021.3083659

Fakhreddine, A., Raffelsberger, C., Sende, M., & Bettstetter, C. (2022). Experiments on Drone-to-Drone Communication with Wi-Fi, LTE-A, and 5G. *Proceedings of the Globecom Workshops (GC Wkshps)*, 904–909. https://doi.org/10.1109/GCWkshps56602.2022.10008743

Fanta-Jende, P., Steininger, D., Kern, A., Widhalm, V., Apud Baca, J. G., Hofsta¨tter, M., Simon, J., Bruckmu¨ller, F., & Sulzbachner, C. (2023). Semantic Real-Time Mapping with UAVs. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, *91*, 157–170. https://doi.org/10.1007/s41064-023-00242-2

Guo, F., Wei, M., Ye, M., Li, J., Mechali, O., & Cao, Y. (2019). An Unmanned Aerial Vehicles Collaborative Searching and Tracking Scheme in Three-Dimension Space. *Proceedings of the Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, 1262–1266. https://doi.org/10.1109/CYBER46603.2019.9066685

Hayat, S., Yanmaz, E., & Bettstetter, C. (2015). Experimental Analysis of Multipoint-to-Point UAV Communications with IEEE 802.11n and 802.11ac. *Proceedings of the Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 1991–1996. https://doi.org/10.1109/PIMRC.2015.7343625

International Federation of Red Cross and Red Crescent Societies. (2022). *World Disasters Report – Trends in Disasters*. Retrieved February 23, 2024, from https://web.archive.org/web/20240223104446/https://www.ifrc.org/sites/default/files/2023-01/20230130_2022_WDR_DataAnnex.pdf

Joint Research Centre. (2023). *The EU 2022 Wildfire Season was the Second Worst on Record*. Retrieved February 23, 2024, from https://web.archive.org/web/20240223104923/https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/eu-2022-wildfire-season-was-second-worst-record-2023-05-02_en

Lauterbach, H., Koch, C. B., Hess, R., Eck, D., Schilling, K., & Nuchter, A. (2019). The Eins3D Project — Instantaneous UAV-Based 3D Mapping for Search and Rescue Applications. *Proceedings of the International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. https://doi.org/10.1109/SSRR.2019.8848972

Lei, L., Tang, A., & Wang, X. (2022). Analysis on the Scalability Issues of Wireless Mesh Networks: Key Factors and Potential Solutions. *Proceedings of the International Conference on Mobile Ad Hoc and Smart Systems (MASS)*, 586–591. https://doi.org/10.1109/MASS56207.2022.00087

Luo, J., Zhao, P., Zheng, F.-C., & Li, L. (2022). Delay Evaluation for Cellular-Connected Drones: Experiments and Analysis. *Proceedings of the Vehicular Technology Conference (VTC-Fall)*. https://doi.org/10.1109/VTC2022-Fall57202.2022.10012989

Mbaitiga, Z., & Shosaku, T. (2022). Assessment of Multi-Drones using City Information for Search and Rescue Operations. *Proceedings of the International Conference on Intelligent Informatics and Biomedical Science (ICIIBMS)*, 379–381. https://doi.org/10.1109/ICIIBMS55689.2022.9971612

Pasini, D., Jiang, C., & Jolly, M.-P. (2022). UAV and UGV Autonomous Cooperation for Wildfire Hotspot Surveillance. *Proceedings of the MIT Undergraduate Research Technology Conference (URTC)*. https://doi.org/10.1109/URTC56832.2022.10002208

Quan, A., Herrmann, C., & Soliman, H. (2019). Project Vulture: A Prototype for Using Drones in Search and Rescue Operations. *Proceedings of the International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 619–624. https://doi.org/10.1109/DCOSS.2019.00113

Scherer, J., Yahyanejad, S., Hayat, S., Yanmaz, E., Andre, T., Khan, A., Vukadinovic, V., Bettstetter, C., Hellwagner, H., & Rinner, B. (2015). An Autonomous Multi-UAV System for Search and Rescue. *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use (DroNet)*, 33–38. https://doi.org/10.1145/2750675.2750683

Schilling, F., Soria, E., & Floreano, D. (2022). On the Scalability of Vision-Based Drone Swarms in the Presence of Occlusions. *IEEE Access*, *10*, 28133–28146. https://doi.org/10.1109/ACCESS.2022.3158758

Seiber, C., Nowlin, D., Landowski, B., & Tolentino, M. E. (2018). Tracking Hazardous Aerial Plumes using IoT-Enabled Drone Swarms. *Proceedings of the World Forum on Internet of Things (WF-IoT)*, 377–382. https://doi.org/10.1109/WF-IoT.2018.8355118

Sende, M., Schranz, M., Prato, G., Brosse, E., Morando, O., & Umlauft, M. (2021). Engineering Swarms of Cyber-Physical Systems with the CPSwarm Workbench. *Journal of Intelligent & Robotic Systems*, *102*(83). https://doi.org/10.1007/s10846-021-01430-1

Steinha¨usler, F., & Georgiou, H. V. (2022). Detection of Victims with UAVs during Wide Area Search and Rescue Operations. *Proceedings of the International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 14–19. https://doi.org/10.1109/SSRR56537.2022.10018756

Terzi, M., Anastasiou, A., Kolios, P., Panayiotou, C., & Theocharides, T. (2019). SWIFTERS: A Multi-UAV Platform for Disaster Management. *Proceedings of the International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. https://doi.org/10.1109/ict-dm47966.2019.9032923

Tinetti, F. G., & Riveros, O. C. V. (2021). Combining Technologies for Aiding Search Missions with Drones. *Proceedings of the International Conference on Computational Science and Computational Intelligence (CSCI)*, 1617–1622. https://doi.org/10.1109/CSCI54926.2021.00013

United Nations Environment Programme. (2022). *Spreading like Wildfire: The Rising Threat of Extraordinary Landscape Fires*. Retrieved February 23, 2024, from https://web.archive.org/web/20240223105515/https://www.unep.org/resources/report/spreading-wildfire-rising-threat-extraordinary-landscape-fires

Viseras, A., Marchal, J., Schaab, M., Pages, J., & Estivill, L. (2019). Wildfire Monitoring and Hotspots Detection with Aerial Robots: Measurement Campaign and First Results. *Proceedings of the International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 102–103. https://doi.org/10.1109/SSRR.2019.8848961

Chatbot Based Public Sensoring to improve Situational Awareness

Hans Betke
Fraunhofer FOKUS
hans.betke@fokus.fraunhofer.de

Sascha Peitzsch
Fraunhofer FOKUS
sascha.peitzsch@fokus.fraunhofer.de

Johannes Boldt
Martin Luther University Halle-Wittenberg
johannes.boldt@wiwi.uni-halle.de

Dustin Reimann
Fraunhofer FOKUS
dustin.reimann@fokus.fraunhofer.de

Thomas Kox
Weizenbaum Institute
thomas.kox@weizenbaum-institut.de

ABSTRACT

In disaster management, any relevant information about the situation can help decision makers to take the right decisions and initiate appropriate response measures. The digitalization of society offers more and more opportunities to gather information from the public. As instant messengers are one of the most widely used communication applications, they offer a broad user base for the collection of situational information. To minimize the effort involved, chatbots offer a promising option for the automated acquisition of specific information via instant messengers. In this paper we present an information system under development for the collection and reproduction of image information for disaster management. We also present the results of a first UTAUT-based evaluation of the current state of development to get an impression of the usability and perceived usefulness of the approach.

Keywords

Situational Awareness, Situation Picture, Chatbot, Public Sensoring, Crowdsourcing

INTRODUCTION

People share an extremely large amount of information in the event of a disaster/crisis via social media (Fathi and Fiedrich 2023). To have access to this information and to disseminate their own information quickly and directly to citizens, many civil protection authorities (CPAs) also use social media and operate their own social media accounts. In addition, there are innovative endeavors to deal more efficiently with information about future developments by implementing new technologies in acute operations (Schütte and Kox 2021). The large amount of different information and social media platforms means a great deal of effort for CPAs to use them effectively and efficiently. For this reason, scientists have been working for several years on approaches to make social media and their services more usable for disaster management purposes (Reuter and Kaufhold 2018). One challenge for which no approach has yet been widely established in practice is the targeted acquisition of situation information from the population.

So called information crowdsourcing is generally not a new topic. Information crowdsourcing approaches are examples where the citizens act as sensors, providing data to interpret and use by scientists or government agencies. The roles as sensors can be passive, for example, downloading an app that allows data gathering via a smartphone, or involve active participation, for example, taking photos or providing other information to the extent that their role can also take on that of an interpreter. For example, gathering data may entail the participants

performing some interpretation where basic training or guidance may be required. Beyond that, citizens involvement can go so far that the citizens play an active role in the problem definition or become part of the interpretation process (Haklay 2013; Vinnell et al. 2021).

Information crowdsourcing in disaster management offers the chance to gather previously unavailable data on e.g. meteorological phenomena (Kempf 2021) or earthquakes (Steed et al. 2019) and thus greatly add to existing observation capabilities of meteorological, geological or emergency services. In addition, crowdsourcing offers a low-cost approach to produce reliable and rapid results (Kox et al. 2021; Steed et al. 2019) and helping affected populations by providing geographic information for disaster response (Fathi and Fiedrich 2023). However, although promising approaches to targeted information crowdsourcing already exist, they are often limited by the reach of their users. For some approaches, users must have certain knowledge or skills, for example in dealing with geoinformation systems, but above all the approaches are mostly based on special software applications that must be installed separately and are not widely used. This means that much of the potential for obtaining information from the population cannot be utilized. Important aspects of the design of corresponding systems are user friendliness and simplicity of the implementation, as it is expected that most users will most likely not be able to accurately report desired phenomena on a professional scale (Kempf 2021).

One idea is therefore to make better use of systems that already have a large user base and enable information exchange, the social media. Instant messaging services such as WhatsApp, Facebook Messenger and Telegram are among the most frequently used applications worldwide (Kemp 2023). These enable simple and direct bidirectional communication with individuals and groups. However, not all CPAs can cope with the effort involved in communicating with many individual people, especially in highly stressful situations such as a disaster. The use of automation techniques in the form of conversational agents could therefore be an approach to ensure the practical applicability of instant messengers. Conversational agents (or chatbots) automatically respond to user input with suitable text messages and can support them in carrying out tasks or provide information themselves. In the area of crowdsourcing physical activities, there are already initial approaches to support the coordination of spontaneous volunteers in disaster response with chatbots (Gerstmann et al. 2019).

In this paper, we present the first steps in our research in progress on the development of targeted information crowdsourcing systems in disaster management using chatbots. To gain first impressions of the usefulness and accessibility of such systems, we have developed a simple chatbot with which a citizen can transmit photos of the situation taken with their smartphone with some meta information to CPAs. The CPA staff can access the information in the form of a dashboard. We have subjected this demonstrator to an initial expert evaluation based on a use case and present the results here. With this paper, we want to provide both scientists and practitioners with an initial assessment of the suitability of chatbot-based approaches for situational awareness and offer a simple solution for the implementation of a system in practical use. While doing so, we investigate the following research question:

RQ: Can chatbot-based situational awareness systems be a user-friendly addition to information retrieval in disaster management?

In the following sections, we first present related work in the field of chatbot-based approaches in disaster management, describe the implementation and features of the software artefact and show the results of an initial evaluation based on the UTAUT-method (Venkatesh et al., 2003). The paper ends with a conclusion summarizing the results and discussing the next steps of the research project.

RELATED WORK

Chatbots are text-based systems, which can communicate with their users via natural language (Gnewuch, Morana, Maedche 2017). They are particularly interesting due to their ability to integrate with other systems as they are mainly used on social media platforms, but also in other applications such as websites and mobile apps (Brandtzaeg and Følstad 2017). This makes them particularly valuable for use in disaster management, where communication via social media plays an important role (Reuter and Kaufhold, 2018). The application of chatbots for emergency and disaster management is an active field of research. In recent years, many different approaches for various use cases have been developed.

The approaches for chatbots in disaster management can be divided into two groups. The first group is concerned with informing and coordinating the general population. For example, Syed et al. (2020) designed an educational chatbot for small and medium sized companies to assist staff in learning about disaster preparedness, ranging from fires and blackouts to cyberattacks. The Richter chatbot on Facebook Messenger (Raymond 2016) provides users

with earthquake and tsunami preparedness information to help them draw up an emergency plan for their household. The “Flood Ai” chatbot (Sermet 2018) teaches users from the public about floods and provides up-to-date information on water levels. Other important use cases involve providing information and guidance to the general population during an emergency (Crook 2016, Ghosh 2019, Ahmady 2020, Boné 2020, Ohtake 2021, Ouerhani 2020, Ovando-Leon 2022) and allowing affected people to contact and communicate with emergency response forces (Clegg 2017, Crook 2016).

The second group focuses on supporting disaster response personnel and decision makers by e.g., facilitating data organization and access (Tsai 2019, Chen 2019, Tsai 2021a, Tsai 2021b). Another example for this is the coordination of spontaneous volunteers (Gerstmann et al. 2019, Ovando-Leon 2022). Furthermore, chatbots can also be applied for informational crowdsourcing. For instance, the chatbot 911bot, which was in practical use, (Crook 2016) was a Facebook Messenger based chatbot that allowed users to report emergencies to the authorities and provided information and advice to first responders. Additionally, it also allowed to send relevant pictures and videos to the authorities. Dharmapuri Sridhar (2017) investigated crowdsourcing situational information for flooding in Chennai, India and presented the web-based tool RiskMap for collecting and verifying crowdsourced information from social media. Participants can report flood-related information, such as location, flood height, and pictures of the flooding, by contacting chatbots on Twitter, Facebook, and Telegram. Tsai (2020) presented a system for the assessment of school building safety after natural hazards, such as earthquakes and typhoons, and for routine inspections. In this use case, people that are typically not structural engineering experts are assigned the task of assessing possible building damages. Hence, they are notified by a chatbot and guided through the inspection procedure and provided further assistance in the process. Furthermore, a dashboard showing the report results and assessed damages is used. This work is built upon in Kung (2020), where transfer learning is used to automatically extract the relevant information from the building damage reports. The chatbot SOCDA (Ohtake 2021) can be used to collect damage reports during natural disasters and can provide the public with relevant information, for example to aid in evacuations. SOCDA was tested in large-scale exercises in Japan in 2020 and 2021.

Despite the examples mentioned here along the entire disaster management lifecycle, the response to the refugee crisis in Germany has been characterized by the need for additional resources and the cooperation of various authorities and organizations under the leadership of a disaster management authority (Dittmer and Lorenz, 2019). Stieglitz et al. (2022) advise disaster managers not to rely on traditional chatbots, that simply gather information, but to deploy intelligent systems that ask information providers for more background information before passing it on to field teams.

Our work aims to build on these demands and previous chatbot implementations for data gathering during disasters and addresses the need to visualize the gathered data to disaster responders which can then be used to inform disaster response workers decision making processes. In the Research in Progress presented here, we explicitly do not consider any further features for analyzing and evaluating the information collected. In order to transfer the approach to practical use, it is important that the information collected is highly trustworthy. Possibilities for integrating existing methods that address this challenge on an organizational (e.g. Mehta et al., 2017) and technical level (e.g. Poblet et al., 2018) will be investigated in future publications.

SOFTWARE ARTIFACT

Our software demonstrator consists of two essential parts for the two user groups involved: A chatbot, the *BurningBot*, through which information providers from the population can transmit images of a situation including meta-information via the instant messenger Telegram, and a dashboard, called *Spell-Dashboard*, which visualizes this information on a situation map for disaster management. In the following, we explain the structure of the system before describing the features.

For a bot to be able to communicate via the Telegram Bot API, it must be registered with the Telegram internal tool BotFather. In the course of this, a name for the conversation ("spell-dashboard-burning-bot") and a username for linking ("@SpellDashboardBurningBot") are assigned and a token is created, which is required to authorize the bot and send requests to the Telegram Bot API. The bot is still active now and the given name can be used to find the chatbot in the Telegram search and to start a conversation. A *BurningBot* Listener Service, developed with Python, runs in the backend. This establishes a connection to the *BurningBot* via the bot token using the HTTP-based Telegram Bot API. The listener is also responsible for giving the *BurningBot* its functionality. All messages that the *BurningBot* sends to the user are defined in the service, such as requesting the image, location and direction. After all the necessary information has been received by the listener, it is converted into a JSON object together with some meta data (user ID, file ID, date and time) and forwarded to a PostgreSQL database. To enable the dashboard frontend to retrieve the data from the database, there is a *BurningBot* API service developed

with NodeJS. The dashboard is an Angular application (v15) and runs on the Red Hat OpenShift Container Platform. All backend services, including those of the *BurningBot*, are provided via OpenShift. Communication between the frontend and the *BurningBot* API takes place via HTTP requests. End users on the disaster management side access the dashboard via HTTPS. Figure 1 shows the components and information flows of the system.

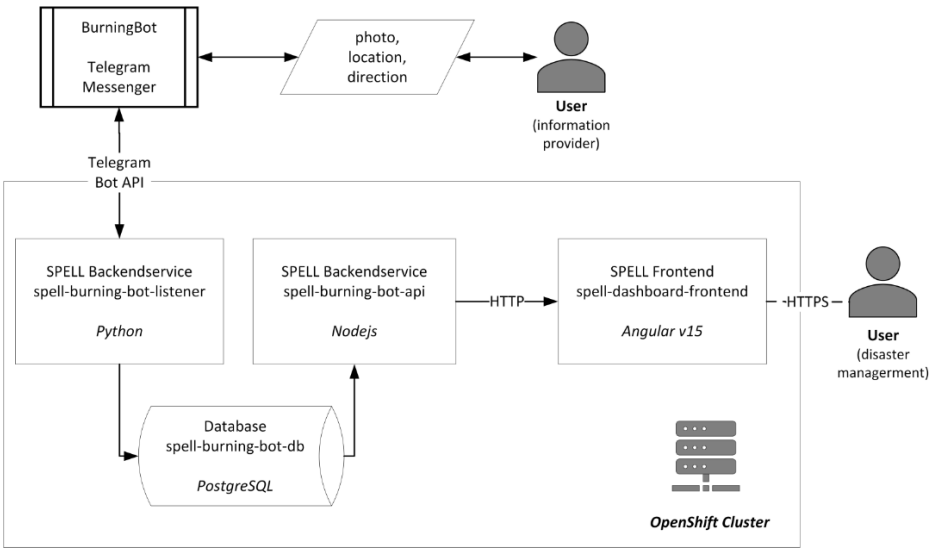


Figure 1. Component diagram of the BurningBot system demonstrator.

The *BurningBot* is a rule-based conversational agent that queries a photo and the corresponding location and direction of the shot in a three-stage conversation. The *BurningBot* first introduces itself and briefly explains how it works and the purpose of information retrieval. It then asks participants to upload an image that will be used as a source of information for CPAs. After the participant has uploaded a picture, the *BurningBot* asks for the participant's location and explains how to share it in the Telegram messenger. If the users have shared their location, the *BurningBot* asks for the cardinal direction in which the picture was taken and provides information on how this can be determined. After entering the cardinal direction, the *BurningBot* thanks the user and the entire conversation can be restarted to share more pictures. An example conversation in German is shown in Figure 2.

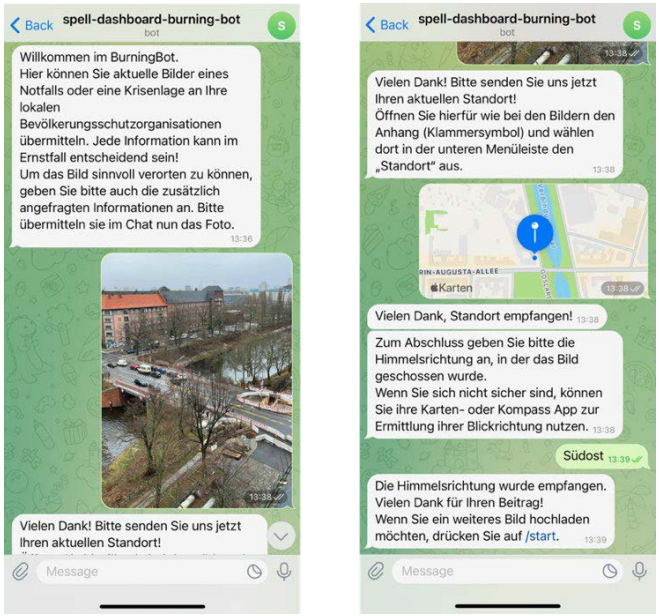


Figure 2. Screenshots of an example interaction with the BurningBot.

The second part of the evaluation involves the visualization of the results on a dashboard that was developed in the context of the SPELL project (https://spell-plattform.de/en/). The participants are provided with the internet address and login information for the dashboard. After login, the participants are presented with a dashboard

landing page, where they can select the "*Lage Berlin*" (translation: "situation picture Berlin") dashboard. They are presented with a panel that shows where participants submitted pictures. A screenshot of an exemplary map is shown in Figure 3.

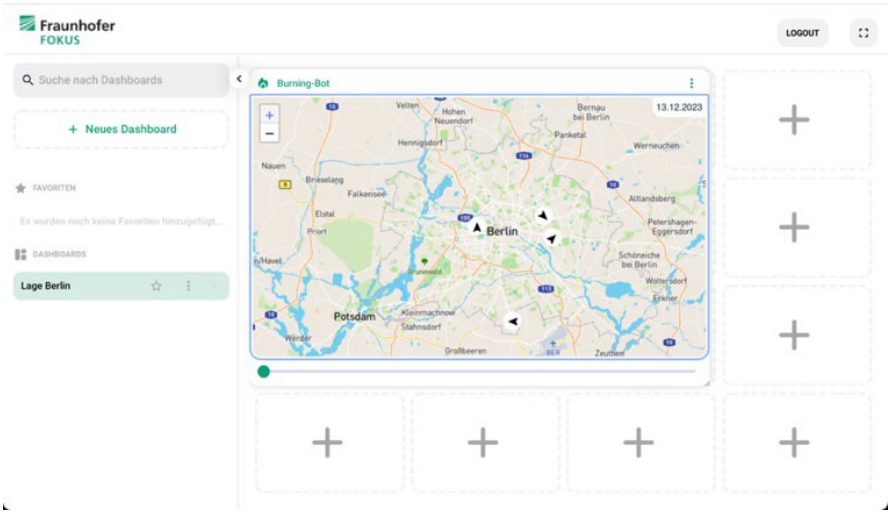


Figure 3. BurningBot panel with map showing locations of submissions.

The slider widget below the map can be used to display the submission status for different times. The positions of the submissions are shown with circle markers, which also indicate the viewing directions of the pictures. Detailed information of the single submissions can be seen by clicking on one of the markers. Figure 4 shows an example submission.

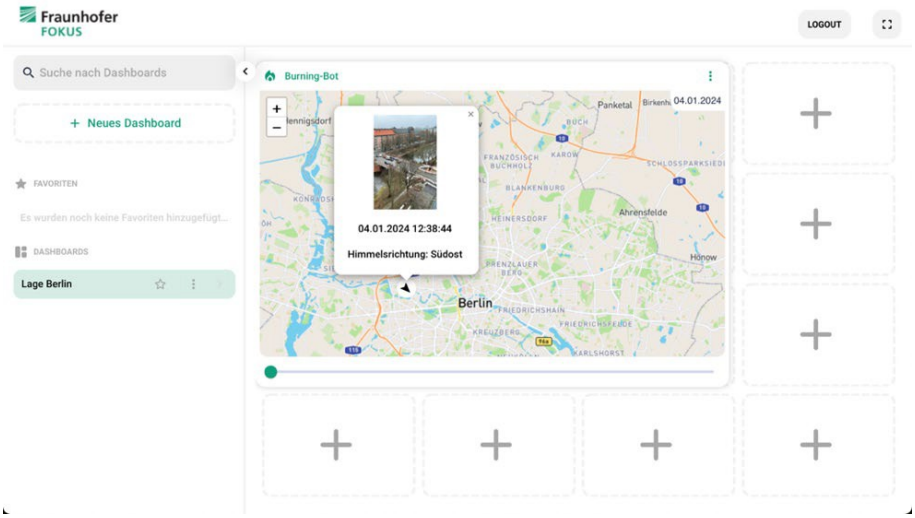


Figure 4. BurningBot panel with detailed information of an example submission.

EVALUATION

To get a first impression of the usability, the perceived usefulness, and user-friendliness of the *BurningBot* and the dashboard system demonstrator, we conducted a UTAUT-based evaluation (Unified Theory of Acceptance and Use of Technology; Venkatesh et al., 2003) of the current state of development. We surveyed people with a background in computer science and with a connection to disaster management as a proxy for the two actual user groups: volunteers, which contribute situational information, and disaster managers, who utilize the submitted information. The questionnaire is based on items and constructs from Venkatesh et al. (2003). The perceived likelihood of adopting the technology is dependent on the effect of four constructs: performance expectancy, effort expectancy, social influence, and facilitating conditions, moderated by age, gender, experience and voluntariness of use (Venkatesh et al., 2003). Certain UTAUT items that do not fit the context of our specific application were not included as they did not fit the purpose of the tool. For example, the item "*If I use the system, I will increase my chances of getting a raise*" was excluded since both the dashboard and the chatbot are not intended to provide economic benefits for the volunteers and disaster managers in our scenario. The items "*It scares me to think that*

I could lose a lot of information using the system by hitting the wrong key" and "*I could complete a job task using the system if I had just the built-in help facility for assistance*" were also omitted. Similarity, the item groups that measure the factors "*social influence*" and "*facilitating conditions*" do not transfer to the scenario at hand and were not used. The questionnaire contained 14 items of the UTAUT, measuring performance expectancy, attitude towards using technology, effort expectancy, self-efficacy, anxiety, and behavioral intent to use the technology. Eight additional items from a questionnaire from Sindhuja p and Dastidar, (2009), which is derived by the Purdue Usability Testing Questionnaire (PUTQ) by Lin et al. (1997) were included to obtain information about user satisfaction and perception of the software design.

The resulting questionnaire contained consequently 22 items for the dashboard and 22 items for the chatbot, which were translated to German. All thematic items consisted of 5-point Likert scales ranging from -2 (strongly disagree) to 2 (strongly agree). In addition, we added demographic items about age and gender. However, we did not include items about the level of education and profession of the participants. The survey was conducted online. Participants could answer items about the chatbot, the dashboard or both systems subsequently. In choosing the sample size, we followed the so-called "10±2 rule" (Hwang and Salvendy, 2010), which states that 8 to 12 respondents are sufficient for evalutation of usefulness of an artifact or technology The experts interviewed all had at least a bachelor's degree, with specialized knowledge in either disaster management, information systems design or both. Finally, 22 questionnaires were completed between December 2023 and January 2024. Ten of the completed questionnaires related to the dashboard while 12 referred to the chatbot. Participants were between 25 and 55 years old. Nine participants identified themselves as female and 13 as male. Due to the preliminary nature of this report, we are only reporting a selection of the data collected, with a focus on the items from UTAUT. All other analyses are currently work in progress.

Performance Expectancy

The performance expectations (Figure 5) for both the bot and the dashboard were generally reported as very high. 50% (dashboard) and 42% (bot) strongly agreed and an additional 40% (dashboard) and 50% (bot) agreed with the statement "*I would find the system useful in my job*" (median value dashboard: 1.5; bot: 1). In addition, 30% (dashboard) and 42% (bot) strongly agreed and additional 30% (dashboard) and 42% (bot) agreed with the statement "*Using the system enables me to accomplish tasks more quickly*" (median value dashboard and bot both: 1).



Figure 5. Survey results for performance expectancy.

Effort Expectancy

Also, effort expectancy (Figure 6) for both the bot and the dashboard were generally reported as very high. 70% (dashboard) and 75% (bot) of the participants strongly agreed and an additional 20% (dashboard) and 25% (bot) agreed with the statement "*It would be easy for me to become skillful at using the system*" (median value dashboard and bot both: 2). 50% (dashboard) and 83% (bot) of the participants strongly agreed and an additional 30% (dashboard) and 8% (bot) agreed with the statement "*I would find the system easy to use*" (median value dashboard:

1.5; bot: 2). 50% (dashboard) and 92% (bot) of the participants strongly agreed and an additional 30% (dashboard) and 8% (bot) agreed with the statement "*Learning to operate the system is easy for me*" (median value dashboard: 1.5; bot: 2).

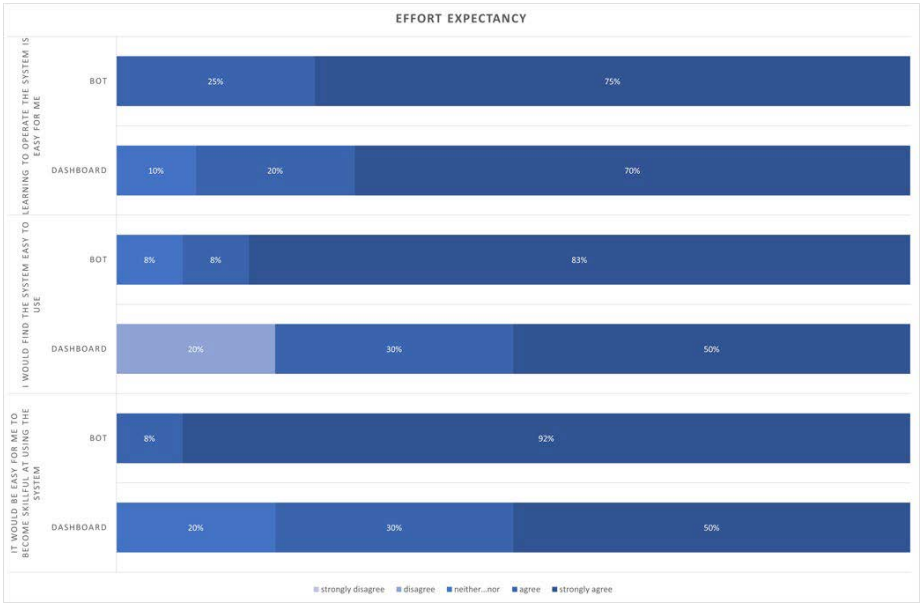


Figure 6. Survey results for effort expectancy.

Attitude Towards Technology

The attitude towards the presented technology (Figure 7) was to some degree indifferent: 50% (dashboard) and 25% (bot) of the participants agreed and an additional 8% (bot) strongly agreed with the statement "*Working with the system is fun*", while 40% (dashboard) and 58% (bot) neither agreed nor disagreed with the statement (median value dashboard: 0; bot: 0.5). However, 20% (dashboard) and 8% (bot) of the participants strongly agreed and an additional 60% (dashboard) and 67% (bot) agreed with the statement "*I like working with the system*" (median value dashboard and bot both: 1). 80% (dashboard) and 83% (bot) of the participants strongly disagreed and an additional 20% (dashboard) and 8% (bot) disagreed with the statement "*Using the system is a bad idea*". 8% of the participants who answered the questionnaire related to the bot, however, agreed with the statement (median value dashboard and bot both: -2).

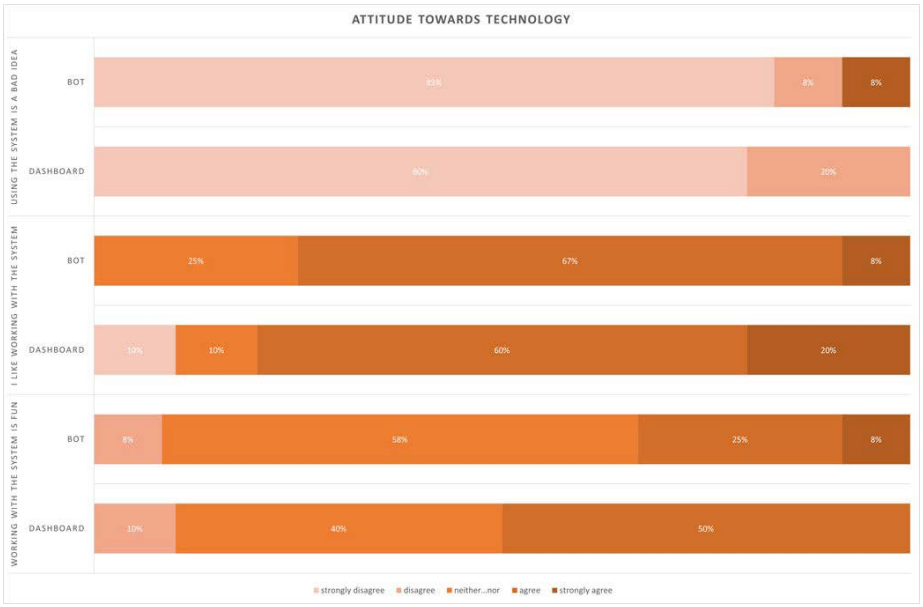


Figure 7. Survey results for attitude towards technology.

Self-Efficacy

Self-efficacy (Figure 8) was measured by asking three questions each starting with "*I could complete a job or task using the system...*" followed by a more specific statement. First, 50% (dashboard) and 67% (bot) of the participants strongly agreed and an additional 50% (dashboard) and 25% (bot) agreed, while 8% of the participants who answered the questionnaire related to the bot disagreed with the statement "*...if there was no one around to tell me what to do as I go*" (median value dashboard: 1.5; bot: 2). Second, 100% (dashboard) and 92% (bot) of the participants strongly agreed and an additional 8% (bot) agreed with the statement "*...if I could call someone for help if I got stuck*" (median value dashboard and bot both: 2). Third, 50% (dashboard) and 67% (bot) of the participants strongly agreed and an additional 30% (dashboard) and 25% (bot) agreed with the statement "*...if I had a lot of time to complete the job for which the software was provided*". However, of the participants who answered the questionnaire related to the dashboard 10% disagreed and an additional 10% strongly disagreed with the statement (median value dashboard: 1.5; bot: 2).

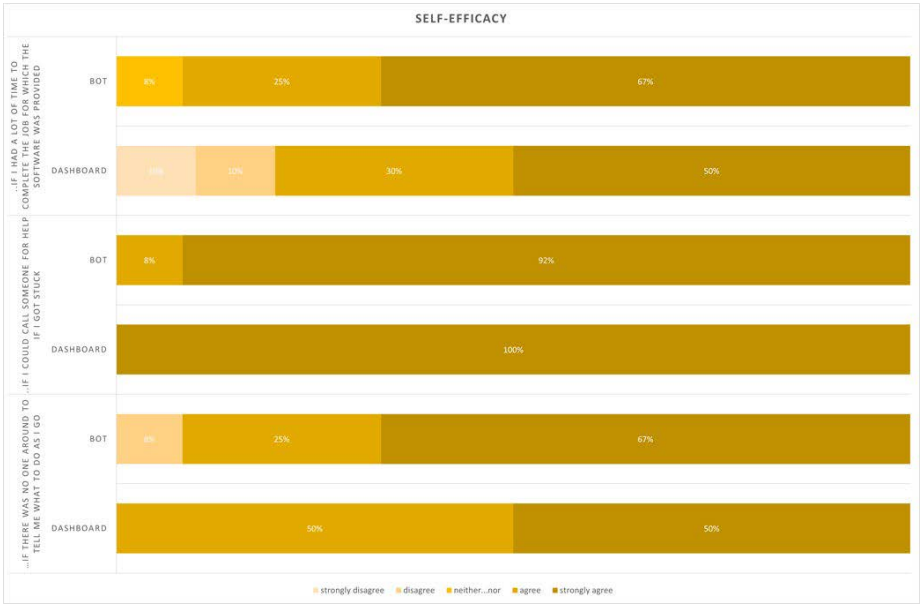


Figure 8. Survey results for self-efficacy.

Behavioral Intention to Use the System

The stated intention to use the bot was indifferent (Figure 10). 42% of the participants neither agreed nor disagreed, while 33% agreed and 17% disagreed with the statement "*I intend to use the system in the next 3 months*" (median value: 0). Regarding the stated intention to use the dashboard, 40% of the participants strongly agreed and an additional 40% agreed with the statement, while 10% each neither agreed nor disagreed or disagreed with the statement (median value bot: 0; dashboard: 1).

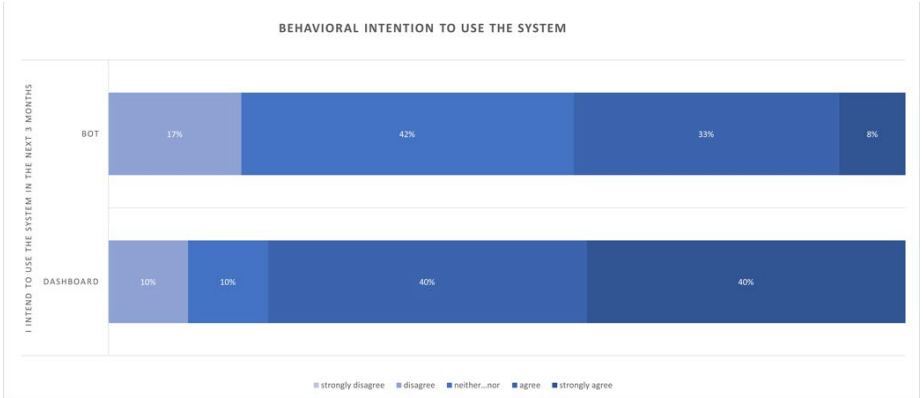


Figure 10. Survey results for behavioral intention to use the system.

Anxiety

Respondents showed no anxiety or concern (Figure 9) about using the dashboard or the bot. 50% (dashboard) and 67% (bot) of the participants strongly disagreed and an additional 40% (dashboard) and 17% (bot) disagreed, while 10% (dashboard) and 17% (bot) neither agreed nor disagreed with the statement "*I feel apprehensive about using the system*" (median value dashboard: -1.5; bot: -2). 50% (dashboard) and 92% (bot) of the participants strongly disagreed and an additional 40% (dashboard) and 8% (bot) disagreed, while 10% of the participants who answered the questionnaire related to the dashboard agreed with the statement "*The system is somewhat intimidating to me*" (median value dashboard: -1.5; bot: -2).

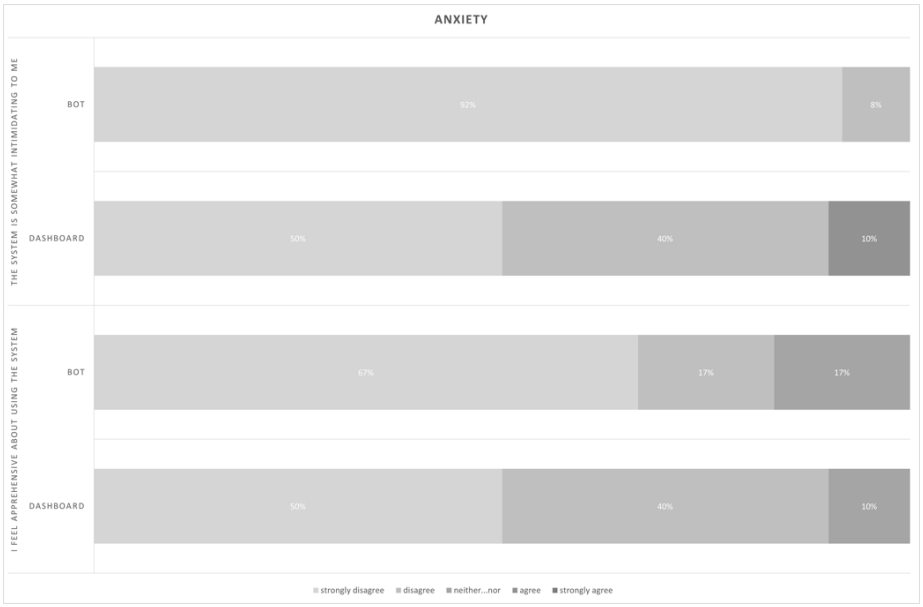


Figure 9. Survey results for anxiety.

Additional Items on Satisfaction and Perception of the Software Design

Regarding the questions on satisfaction with the software design of the dashboard and the bot, we present a selection of the items here (Figure 11). In general, the participants were satisfied with the software design. 40% of the participants strongly agreed and an additional 30% agreed with the statement "*I like the design of the dashboar*d." 10%, however, strongly disagreed with the statement and 20% were indecisive (median value: 1). In addition, 30% of the participants strongly agreed and an additional 30% agreed with the statement "*The dashboard provides clarity and consistency of wording*". Again, 10% strongly disagreed with the statement and 30% were indecisive (median value: 1).

Regarding the color scheme and the arrangement of the dashboard, participants gave positive feedback. First, 40% of the participants strongly agreed and an additional 50% agreed with the statement "*Color codes used in the dashboard are easily distinguishable*" (median value: 1). Second, 50% of the participants strongly agreed and an additional 40% agreed with the statement "G*rouping and ordering of menu options [in the dashboard] are logical*" (median value: 1.5). Third, 30% of the participants strongly agreed and an additional 40% agreed with the statement "*I can find easily what I'm looking for on the dashboard*", while 10% each strongly disagreed, disagreed or were indecisive (median value: 1).

The participants were not sure whether frequent and regular use of the dashboard and the bot was necessary to handle the technology. 20% (dashboard) of the participants strongly agreed and an additional 20% (dashboard) and 17% (bot) agreed with the statement "*The user needs to be a frequent user*", while 20% (dashboard) and 17% (bot) neither agreed nor disagreed with the statement. In contrast, 20% (dashboard) and 33% (bot) of the participants strongly disagreed and an additional 20% (dashboard) and 33% (bot) disagreed with the statement (median value dashboard: 0; bot: -1).

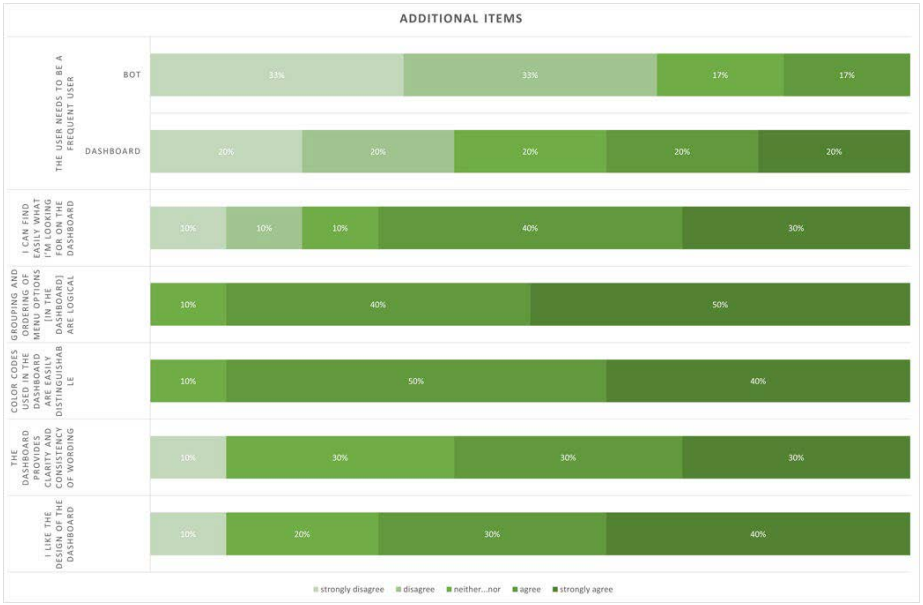


Figure 11. Survey results for additional items.

Limitations

Not just because of the provisional nature of the software, there are several limitations to this study: The evaluation only included part of the evaluation items, was conducted with only a few test persons and not under real-life conditions. It can be assumed that the statements will change under these conditions. The current prototype, for example, does not provide any information on the expected extent of contributions during a real-life crisis event. However, as lay persons networks made up of e.g. voluntary weather observers (Kox et al. 2021), or digital volunteers and professionalized digital volunteers institutionalized in established CPAs (Fathi and Fiedrich 2023) also form a community of practice and their willingness to engage in such a crowdsourcing task should not be underestimated (Kox et al. 2021). Additionally, in a real-life emergency context we would expect a bias between urban over rural areas, as it can be expected that the density of crowd observations will be higher in higher populated urban regions (Meier et al. 2017).

CONCLUSION

In this paper, we used a software demonstrator to investigate how potential users evaluate chatbot-based applications for situational awareness in disaster management. The approach presented here uses citizens and their mobile devices as sensors that interact with the object of interest to some extent, but do not provide any interpretative services. The participants in this use case mainly take the roles as *active sensors* by taking photos and providing context information. Images transmitted via chatbot with location and time stamp are displayed to decision-makers in disaster management in the form of a map-based dashboard. Both frontends were evaluated by expert groups based on established methods for evaluating technology acceptance after they had familiarized themselves with the system in a simple use case. The results showed a generally positive assessment of the user-friendliness and perceived usefulness of the system for both user groups.

Performance expectations and effort expectancy for both the bot and the dashboard were generally reported as very high. The respondents showed no anxiety or concern about using either of these technologies. The software design of the dashboard was generally evaluated positively. However, the stated intention to use the bot was indifferent, suggesting that other factors, which were not covered in this study, still might be affecting this variable. The factors "*social influence*" and "*facilitating conditions*" were omitted in this preliminary study and might be included in future research.

Despite the limitations of this work, such as the limited number of participants in the evaluation in the current research in progress and the fact that the software demonstrator only has basic functionalities, the clearly positive assessment of the system and the idea behind it suggests that corresponding solutions could offer significant practical added value and make further research in this area appear promising. In future research steps, we want to extend the existing demonstrator to investigate chatbot-based querying and dashboard-supported evaluation of different types of information. To better assess the added value and acceptance of such systems in practice, we

are aiming for an evaluation in the context of a larger disaster control exercise with more realistic operating conditions and a larger number of participants. We want to methodically summarize the design knowledge acquired in the course of this work for the implementation of such systems in a design theory to provide scientists and practitioners with a suitable starting point for the implementation of their own systems. Another research question that has been investigated for several years in the field of information gathering via social media concerns the IT-supported evaluation of the information collected. In addition, using sophisticated data analysis tools for automatic classification the crowdsourced information could help to build up databases of disaster impact that could not otherwise be created.

ACKNOWLEDGMENTS

This work is funded by the German Federal Ministry for Economic Affairs and Climate Action under grant No. FKZ 01MK21005B SPELL. The contribution of Thomas Kox was part of the joint project "Weizenbaum Institute for the Networked Society" (FKZ 16DII131), funded by the German Federal Ministry of Education and Research. We thank all people who participated in the evaluation of the dashboard and the chatbot.

REFERENCES

Ahmady, S. E., and Uchida, O. (2020). Telegram-based chatbot application for foreign people in Japan to share disaster-related information in real-time. In *2020 5th International Conference on Computer and Communication Systems* (ICCCS)(pp. 177-181). IEEE.

Boné, J., Ferreira, J. C., Ribeiro, R., and Cadete, G. (2020). Disbot: A Portuguese disaster support dynamic knowledge chatbot. In *Applied Sciences*, 10(24), 9082.

Brandtzaeg, P.B., Følstad, A. (2017). Why People Use Chatbots. In Kompatsiaris, I., et al. *Internet Science. INSCI 2017. Lecture Notes in Computer Science*, vol 10673. Springer, Cham. https://doi.org/10.1007/978-3-319-70284-1_30

Chen, J. Y., Tsai, M., Yang, C., Chan, H., and Kang, S. (2019). Chatbot System for Data Management: A Case Study of Disaster-related Data. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction* (Vol. 36, pp. 306-309). IAARC Publications.

Clegg, J. (2017). Rescue. io: A chatbot solution for emergency situations. In *Square One Labs* https://medium.com/square-one-labs/rescue-io-a-chatbot-solution-toemergency-situations-dd267f174554.

Crook, J. (2016). *911bot is a chat bot that could save your life*. https://techcrunch.com/2016/05/08/911bot-is-a-chat-bot-that-could-save-your-life

Dittmer, C. and Lorenz, D. (2019). Disaster situation and humanitarian emergency – in-between responses to the refugee crisis in Germany. *International Migration*, 59(3), 96-112. https://doi.org/10.1111/imig.12679

Dharmapuri Sridhar, M. P. (2017). *Real-time flood mapping for disaster management decision support in Chennai* (Doctoral dissertation, Massachusetts Institute of Technology).

Fathi, R., and Fiedrich, F. (2024). Digital Volunteers in Disaster Management. In D. Burghardt, E. Demidova, and D. A. Keim (Eds.), *Volunteered Geographic Information: Interpretation, Visualization and Social Context* (pp. 265–276). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-35374-1_13.

Gerstmann, S., Betke, H., and Sackmann, S. (2019). Towards Automated Individual Communication for Coordination of Spontaneous Volunteers. In *Proceedings of the 16th Annual Global Conference on Information Systems for Crisis Response and Management (ISCRAM 2019), Spain*.

Ghosh, P., Raihan, M., Islam, M. T., and Rahaman, M. E. (2019). Safeguard: A prototype of an application programming interface to save the disaster affected people. In *2019 10th International Conference on Computing, Communication and Networking Technologies* (ICCCNT) (pp. 1-5). IEEE.

Gnewuch, U., Morana, S. and Maedche, A. 2017. *Towards Designing Cooperative and Social Conversational Agents for Customer Service*.

Haklay, M. (2013). Citizen science and volunteered geographic information: Overview and typology of participation. In D. Sui, S. Elwood, & M. Goodchild (Eds.), Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice (Vol. 9789400745, pp. 1–396). Springer. https://doi.org/10.1007/978-94-007-4587-2

Hwang, W. and Salvendy, G. (2010) Number of people required for usability evaluation, *Communications of the ACM*, 53, 5, 130–133.

Kemp, S. (2023). *DIGITAL 2023: Global Overview Report*, https://datareportal.com/reports/digital-2023-global-overview-report

Kempf, H. (2021). Experience from large-scale crowdsourcing via weather apps. In *Australasian Journal of Disaster and Trauma Studies*, 25(3), 87–93.

Kox, R., Wentzel, G., Böttcher, L., and Freundl, G. (2021). Build and Measure: Students report weather impacts and collect weather data using self-built weather stations. In *Australasian Journal of Disaster and Trauma Studies*, 25(3), 79–86.

Kung, H. K., Hsieh, C. M., Ho, C. Y., Tsai, Y. C., Chan, H. Y., and Tsai, M. H. (2020). Data-augmented hybrid named entity recognition for disaster management by transfer learning. In *Applied Sciences*, 10(12), 4234.

Lin, H. X., Choong, Y. Y., and Salvendy, G. (1997). A proposed index of usability: A method for comparing the relative usability of different software systems. In *Behaviour & Information Technology*, *16*(4–5), 267–277.

Mehta, A. M., Bruns, A., & Newton, J. (2017). Trust, but verify: social media models for disaster management. *Disasters*, 41(3), 549-565.

Meier, F., Fenner, D., Grassmann, T., Otto, M., and Scherer, D. (2017). Crowdsourcing air temperature from citizen weather stations for urban climate research. In *Urban Climate*, 19, 170–191. https://doi.org/10.1016/j.uclim.2017.01.006

Ohtake, K. (2021). Reseach and Development on Technologies for Real-time Analysis of Social Wisdom. In *NICT NEWS 2021*, 3, 8-9.

Ouerhani, N., Maalel, A., and Ben Ghézela, H. (2020). SPeCECA: a smart pervasive chatbot for emergency case assistance based on cloud computing. In *Cluster Computing*, 23, 2471-2482.

Ovando-Leon, G., Veas-Castillo, L., Gil-Costa, V., and Marin, M. (2022). Bot-Based Emergency Software Applications for Natural Disaster Situations. In *Future Internet*, 14(3), 81.

Poblet, M., García-Cuesta, E., & Casanovas, P. (2018). Crowdsourcing roles, methods and tools for data-intensive disaster management. *Information Systems Frontiers*, 20(6), 1363-1379.

Raymond, P. (2016). *Meet Richter: Earthquake Preparation Bot*. https://chatbotslife.com/meet-richter-earthquake-preparation-made-a-little-simpler-d143128a544

Reuter, C., and Kaufhold, M. A. (2018). Fifteen years of social media in emergencies: a retrospective review and future directions for crisis informatics. In *Journal of Contingencies and Crisis Management*, 26(1), 41-57.

Schütte, P., and Kox, T. (2022). Vor die Lage—Jetzt und morgen. Herausforderungen von BOS im Umgang mit neuen Technologien und Digitalisierung. In Fekete, A. (Ed) *Kritische Infrastruktur und Versorgung der Bevölkerung* (pp. 19–22). Springer International Publishing.

Sermet, Y., and Demir, I. (2018). An intelligent system on knowledge generation and communication about flooding. In *Environmental modelling & software*, 108, 51-60.

Sindhuja p, P. N., and Dastidar, S. G. (2009). Impact of the Factors Influencing Website Usability on User Satisfaction. In *Journal of Management Research*, *8*(12), 54-66.

Syed, H. A., Schorch, M., and Pipek, V. (2020). Disaster learning aid: A chatbot centric approach for improved organizational disaster resilience. In *Information Systems for Crisis Response and Management Conference; Virginia Tech: Blacksburg, VA, USA* (pp. 448-457).

Steed, R. J., Fuenzalida, A., Bossu, R., Bondár, I., Heinloo, A., Dupont, A., Saul, J., and Strollo, A. (2019). Crowdsourcing triggers rapid, reliable earthquake locations. In *Science Advances*, 5(4), eaau9824. https://doi.org/10.1126/sciadv.aau9824

Stieglitz, S., Hofeditz, L., Brünker, F., Ehnis, C., Mirbabaie, M., and Ross, B. (2022). Design principles for conversational agents to support Emergency Management Agencies. In *International Journal of Information Management*, *63*, 102469.

Tsai, M. H., Chan, H. Y., Chan, Y. L., Shen, H. K., Lin, P. Y., and Hsu, C. W. (2021a). A chatbot system to support mine safety procedures during natural disasters. In *Sustainability*, 13(2), 654.

Tsai, M. H., Yang, C. H., Chen, J. Y., and Kang, S. C. (2021b). Four-stage framework for implementing a chatbot system in disaster emergency operation data management: A flood disaster management case study. In *KSCE Journal of Civil Engineering*, 25(2), 503-515.

Tsai, M. H., Chan, H. Y., and Liu, L. Y. (2020). Conversation-based school building inspection support system. In *Applied Sciences*, 10(11), 3739.

Tsai, M. H., Chen, J. Y., and Kang, S. C. (2019). Ask Diana: A keyword-based chatbot system for water-related disaster management. In *Water*, 11(2), 234.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. In *MIS Quarterly*, *27*(3), 425-478. https://doi.org/10.2307/30036540

Vinnell, L. J., Becker, J. S., Scolobig, A., Johnston, D. M., Marion L. Tan, and McLaren, L. (2021). Citizen science initiatives in high-impact weather and disaster risk reduction. In *Australasian Journal of Disaster and Trauma Studies*, 25(3), 55–60.

Modular Multi-Platform Interface to Enhance the Situational Awareness of the First Responders

Xabier Oregui*
Vicomtech Foundation Basque Research and Technology Alliance (BRTA), 20000 Donostia, Spain;
xoregui@vicomtech.org

Izar Azpiroz
Vicomtech Foundation Basque Research and Technology Alliance (BRTA), 20000 Donostia, Spain;
iazpiroz@vicomtech.org

Verónica Ruiz
Universidad Politécnica de Madrid, Madrid, 28040, Spain;
veronica.ruiz@upm.es

Blanca Larraga
Universidad Politécnica de Madrid, Madrid, 28040, Spain;
blanca.larraga@upm.es

Álvaro Gutiérrez
Universidad Politécnica de Madrid, Madrid, 28040, Spain;
a.gutierrez@upm.es

Igor G. Olaizola
Vicomtech Foundation Basque Research and Technology Alliance (BRTA), 20000 Donostia, Spain;
iolaizola@vicomtech.org

ABSTRACT

Augmented Reality technologies can provide useful information to First Responders in rescue scenarios if their design is properly adapted to their cognitive and circumstantial requirements. Indeed, biosignals and environmental sensors can improve the situational awareness of the professionals working in emergency situations, supporting the decision-making process in critical moments. This study presents the architecture to create a Multi-Platform Interface to Enhance the Situational Awareness of First Responders that can be customized through an external dashboard. The key to individualizing the interface is the flexible configuration of sensor modules, normally manipulated by the Command Center according to the First Responders' rescue context.

Keywords

Augmented Reality, Cognitive Load, Situational Awareness, Angular, MQTT

INTRODUCTION

First Responders (FRs) operate in logistically demanding, mentally challenging, and time-sensitive circumstances. This is why, over the last few decades, there has been a growing interest in the development of technology that facilitates their rescue tasks. Augmented reality technologies, for instance, have been developed to provide important information to officers operating in catastrophic situations, as reflected in studies such as Arregui et al., 2022; Kapalo et al., 2018; Neustaedter et al., 2019. In this context, the International Forum to Advance First Responder Innovation (IFAFRI, 2023) has detected 10 FR common global capability gaps, a list that facilitates the characteristics of the innovative technologies to augment FR safety, efficiency, and effectiveness. Between those gaps, real-time information-sharing seems primordial to track the location of FRs, and monitor and analyze threats and hazards.

*xoregui@vicomtech.org

The incorporation of information available from multiple and nontraditional sources has been detected as an important capability gap, including the option to monitor physiological signs of FRs. Apart from sensor-based information, interoperable communications with FRs are supposed to enhance FRs safety and efficiency independent of environmental conditions.

Considering this circumstance, according to Arregui et al., 2022; Kapalo et al., 2018 appropriately designed Augmented Reality (AR) tools provide essential information to FRs to support decision-making and improve their Situational Awareness (SA) On top of this, The organization of the data in modules makes its processing more efficient. Indeed, the comprehension of the information processed by FRs in catastrophic circumstances varies according to their cognitive load as indicated by Meneguzzi et al., 2012. Despite enhancing the abilities of FRs, Gutie´rrez et al., 2023 highlights that the volume and manner in which these details are presented can impact their cognitive workload. The modulation of the amount of information displayed in AR interfaces facilitates focusing on the main goal of technology incorporation in emergency cases: the FRs' safety and minimizing risk exposures. The tracking and monitoring of biosignals allow for estimating the cognitive load and, therefore, adapting information exposure as already studied by Frederick et al., 2010; Kahneman and Beatty, 1966.

As identified by IFAFRI, 2023, another essential factor in catastrophic situations is the transmission of instantaneous information between agents present at the scene and the control center. The relevancy of this communication is reflected in the effort programmed by Canada in order to include the sharing of photos and videos between callers and 9-1-1 call centers within Next Generation 9-1-1 programs summarized in Camp et al., 2000; NG911, 2023. Ludwig et al., 2013; Neustaedter et al., 2019 indicate that real-time data exchange is equally significant between the RFs and the control center in emergency scenarios. In this way, the data collected by the sensors and devices of the FRs provides strategic advice in the control centers, as reminded in Ludwig et al., 2013.

Park et al., 2014 insists on the need for ergonomic tools that fit the usual equipment of FRs as a prerequisite for providing maximum efficiency and minimizing the difficulty of their task. A modality that facilitates the ergonomics of the FRs and the transmission of information between operators is voice communication, as remarked by Choong and Salvendy, 2021. When one of the essential tools is in the hands of the user in augmented or virtual reality situations, Monteiro et al., 2021 outlines the efficiency of voice as one of the most efficient interaction means. In this case, the use of voice can provide a double contribution: communication with external operators, and the ability to modulate the display, ensuring a completely hands-free context for the FRs. In this context, recent research as Arregui et al., 2022 shows the benefits of prototypes created according to these circumstances, such as the Smart Helmet presented by Ferna´ndez Garc´ıa et al., 2023. This tool combines sensors, artificial intelligence, AR, and personal protection to increase the situational awareness of FRs.

Within this context, this study aims to introduce an AR interface design characterized by the following objectives: modularity to adjust information exposure according to FRs' cognitive condition; ergonomic or hand-free adaptability; architecture flexibility to enable simultaneous deployment in interconnected distinct devices; and real-time data actualization capability. This presentation includes, on one hand, the description of the main architecture that provides a modular multi-device display, and it details the combination of operational modules connected through an MQTT-based communication protocol. On the other hand, it summarizes the generated prototype that demonstrates the versatility of the generated AR interface. This part includes its deployment inside a display added to an adapted FR helmet and the simultaneous visualization in the control center, which adjusts through a dashboard the tools to be displayed. The last section summarizes the impact of the experience of the pilots on the AR interface design, and finishes with the pending and future work perspectives.

DESCRIPTION OF THE SYSTEM

The Augmented Reality (AR) interface presented in this work, The Modular Multi-Platform Interface (MMPI), proposes a technology to aggregate and display essential information in order to enhance situational awareness and support decision-making. This mechanism improves the transfer of information between the agents in the emergency field and the technicians in the control center. All while keeping from overloading the users with information and preventing mistakes as a result of this extra-information exposure. This data comprises essential details such as the location and orientation direction of First Responders (FRs), the positioning of other FR team members, and noteworthy points of interest—such as targets or alerts regarding potential dangers or events tagged by other FRs or directly from the Command Center.

The design of the AR display is characterized according to the needs of the FRs in the emergency field since the main objective is to extend their capacities when working in crisis zones. Due to ergonomic reasons, the AR is engineered to be projected in the one-eyed transparent AR Head-Up Display (HUD) integrated into a standard helmet enhanced with a processing unit and camera, called Smart Helmet on the Ferna´ndez Garc´ıa et al., 2023;

but it is similarly displayable on an opaque screen of a mobile phone or a tablet. This is a result of the flexible design based on the dynamic web application framework Angular detailed by Jain et al., 2014, technology that facilitates the adjustment and duplicated reproduction of the interface on different displays. In this manner, all biosignal, sensor, and positional information displayed on the interface is available simultaneously to the rest of the FRs and the Command Center.

The procedure for transferring information across multiple devices is based on Message Queuing Telemetry Transport (MQTT), a lightweight messaging protocol designed for use in low-bandwidth, high-latency, or unreliable networks. The biosignal, contextual, and positional data recorded by distinct sensors or indicated by the FRs is published through previously defined topics in the MQTT broker. The subscription to those channels enables the real-time update of the information projected in the AR.

On the other hand, the FRs in emergency fields operate in dynamic and demanding operational scenarios, and therefore, it is crucial to equip the FR with comprehensive and adapted information facilitating real-time decision-making instead of over-saturating and complicating the situational awareness and decision-making processes. That is accomplished by adapting the modular configuration of the AR display through a dashboard manipulated in the control center. The cognitive load is the key indicator to avoid information saturation or overloading and provide an assumable data quantity. Gutie´rrez et al., 2023 provide measures to approximate it.

The following Global Architecture section includes all the technical requirements involved in the simultaneous deployment of the AR interface in distinct devices, which are distributed in two different physical environments: the Emergency Field and the Control Center.

Global Architecture

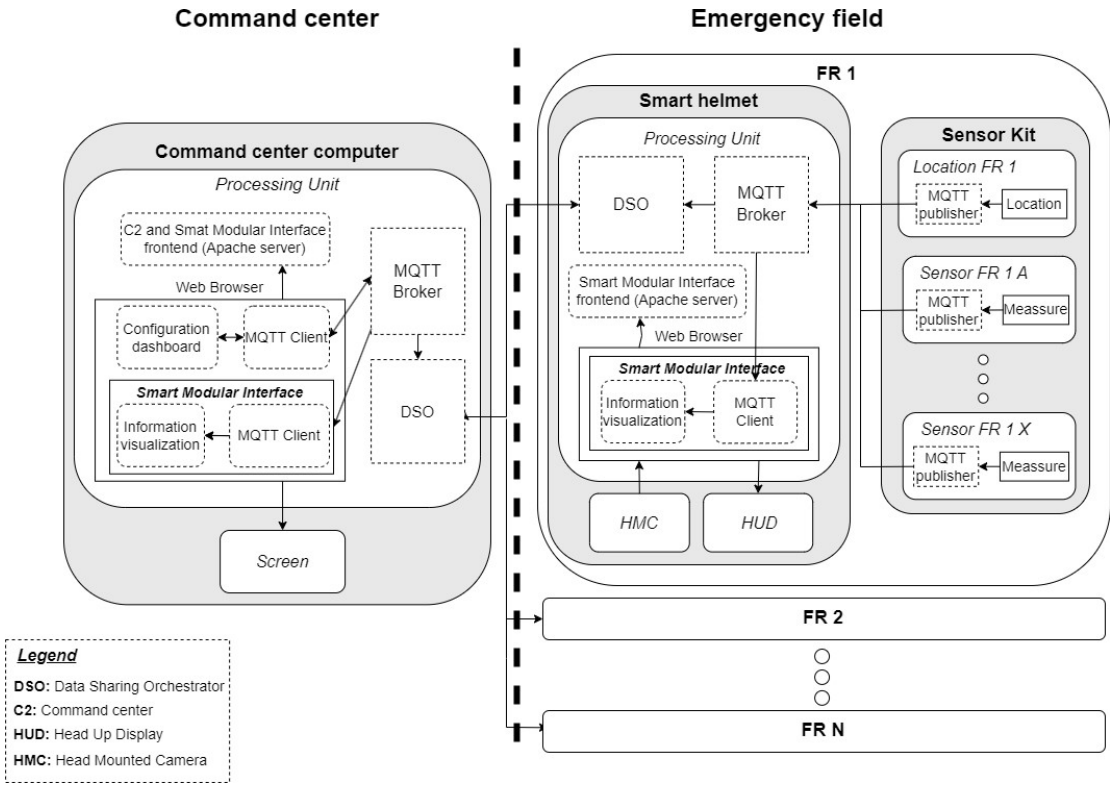


Figure 1. Use case architecture

The Modular Multi-Platform Interface (MMPI), an AR interface, has been designed to perform in a highly adaptable architecture resulting from different emergency scenarios, with changeable FR and Command Center (C2) organization. The presented architecture facilitates the intercommunication between multiple operational nodes in order to retrieve, process, and display information relevant to the FR. The architecture depicted in Figure 1 represents the main technology division that occurs in emergency response services, the digital frameworks corresponding to the Emergency Field and the Control Center. In the Emergency Field, the FRs illustrate the necessary digital configuration to display the AR in each deployed field operator. At the same time, the Central

Operation Computer symbolizes the Control Center and the required technology to overview the information of the FRs and the dashboard to customize their MMPI as depicted in Figure 4.

Both sides of the use case share many components. On the one hand, both contain a processing unit (Laptop, Raspberry, ...) that includes an MMPI, an MQTT broker, and a Data Sharing Orchestrator (DSO). The MQTT broker is responsible for collecting and managing the data flowing within the network established by the FRs and the control center. The MMPI consumes the messages available in the MQTT broker topics to supply the FR display or the control center screen with real-time data. Inside the processing unit, another common component is the Data Sharing Orchestrator. The DSO is a module that fulfills two tasks: forwarding the information from the sensors of each FR to the rest of the MQTT Brokers and filtering the data of other FRs based on the user's cognitive load to avoid overloading Gutie´rrez et al., 2023.

On the other hand, the control center and the emergency equipment differ in important characteristics. The Control Center includes a screen, in replacement of an AR device, and a Configuration Dashboard (CD) that personalizes the modular appearance MMPI of each FR. In addition, the CD can provide externally defined coordinates of points of interest in the Emergency field. Since the Smart Helmet does not offer the FR ways to interact with the MMPI, that is the unique manner to update information not registered by sensors. On the other hand, an FR in the Emergency Field carries, apart from a Smart Helmet that includes the Head Mounted Camera and Head Up Display, the necessary rescue equipment. This equipment contains an individualized collection of sensors, denoted as the Sensor Kit in this context. Indeed, the following list highlights the instruments considered within the RescueR project as part of the Sensor Kit:

* Smartwatches collecting the wearer's physiological signals
* Wireless signal detectors encompassing Bluetooth and WiFi frequencies, enabling potential victim localization in obstructed environments (e.g., under debris).
* Vital sign detectors designed to discern pulsation through solid barriers, facilitating life detection in scenarios where subjects are obscured by walls or similar impediments.
* Chemical sensors engineered to identify hazardous airborne constituents, such as methane or diminished oxygen levels, contributing to environmental safety monitoring.
* Localization devices employing GPS Galileo for outdoor positioning and leveraging inertial sensors, specifically mobile phone gyroscopes, for indoor localization purposes.

In order to ensure the transference of the data recorded by these sensors to the AR interface, the MMPI in this case, is necessary to publish those values in a previously agreed message format and topic for the MQTT. These topics shape the modules that are visualized in the MMPI, thus characterizing the modularity of the AR interface.

The Modular Multi-Platform Interface, the main contribution of this paper, is the technological cornerstone that facilitates the integration of all data gathered from different Sensor Kits. Its modularity, according to FR's cognitive load and adaptability to distinct devices, is characterized by its Angular-based front-end development. The main advantage of using Angular is its flexibility and structure that is specifically designed to generate individual components that can be grouped, nested or even replicated and reused in multiple sites. In addition, it is a widely used framework on which multiple libraries have been created (including a dependency inspector to facilitate their integration) that facilitates the development of more complex solutions. These characteristics fit the objectives of AR interface design, as mentioned above, a generation of an extensible and adaptable interface depending on the equipment assigned to each FR. Deployment options permit its installation on either an Apache server within the local device interfacing with the display or on a server interconnected within the shared network.

In addition to this, like many other JavaScript frameworks, its results are accessible from any modern web browser offering a lot of flexibility in terms of where to display the information. In this case, being an interface designed specifically for AR displays, the background of the application by default is a black background. Normally, AR displays interpret the black color as a transparent pixel allowing to combine reality with an artificial scenario. When the interface needs to show more information, such as a camera feed, the AR display can function as a regular opaque screen. Nonetheless, its functionality relies upon the interconnection between the display and the processing units, alongside the camera requisites delivering the camera feed.

Prototype

This section describes the different elements that conform complete architecture of the use case either on the Control Center side or in the emergency field side: The MMPI, the Configuration Dashboard and the Smart Helmet.

Modular Multi-Platform Interface (MMPI)

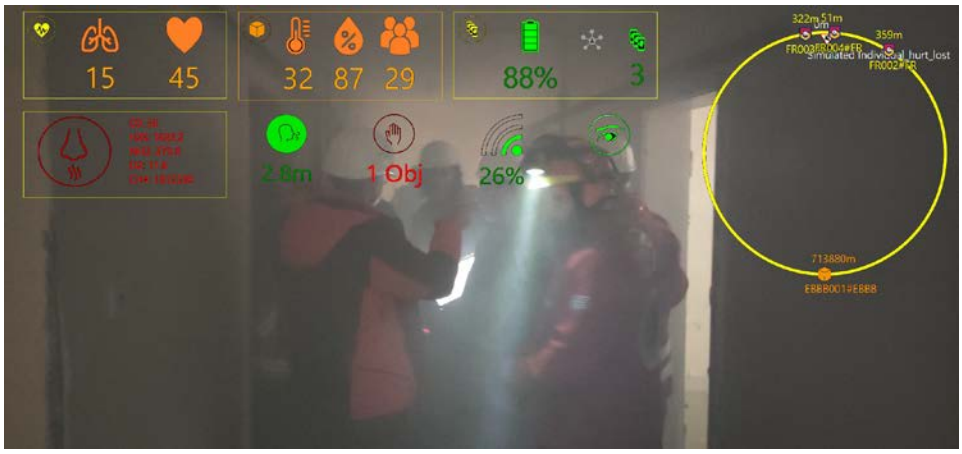


Figure 2. Example of the Modular Multi-Platform Interface (MMPI) visualization in a FR helmet HUD

The MMPI, has been generated to test its adaptability across the display in different devices, the implementation of the distinct-format data, as well as the connectivity of the components responsible for data management. As summarized in the Methodology section, the same AR interface is deployed on distinct displays: a laptop and a Head-Up Display (HUD) integrated into an FR Helmet, representing the communication segmentation of catastrophic circumstances. This duplication facilitates direct transmission of the information gathered by the FRs in the Emergency Field. The AR interface depicted in Figure 2 exhibits crucial information for First Responders (FRs) to navigate crisis scenarios effectively. Within this display, two distinct categories emerge: the passive data and the Tool-associated Information. The former encompasses data that is directly transmitted to the FR without requiring his interaction: for example, biosignals, location, or network status. Conversely, the Tool associated Information section includes the output of devices manipulated by the FRs during the rescue operation; for example such as radars or gas detectors (see Table 1 for further description of some examples applied during validation). The segmentation of all this information in modules facilitates a comprehensive overview of both passive and tool-related data, empowering FRs to make informed decisions while engaging with diverse tools and technologies. These Situational Awareness enhancing Modules and corresponding sensors are detailed in Table 1.

The information distribution on the screen is important to focus the attention of the user according to contextual needs. The focus of a First Responder (FR), for example, typically lies in the lower field of vision as they concentrate on locating victims. Therefore, in this case, the AR layer is required to be positioned at the top of the field of view to ensure the lower part remains unobstructed, as visualized in Figure 2.

During concrete moments in rescue operations, the FRs necessitate a distinct kind of information visualization, a configuration not summarizable with small icons and values in a small space fraction of the AR interface. Instead, this critical information needs to be displayed over the complete AR scene or screen. Therefore, two opaque interfaces have been included, display modes changeable from the Control Center. The two non-AR modes integrated are the Camera Feed mode, which allows the direct reproduction of camera output on the FR display (useful for infrared cameras or IA-generated image processing, Ferna´ndez Garc´ıa et al., 2023) or full-screen map mode that illustrates the position of the FRs and points of interest, as depicted in Figure 3.

Configuration Dashboard

This web tool has been designed to enable the Control Center to modify the depiction of modules for a specific FR and to facilitate the transmission of information updates necessary to keep the emergency field agent on track of their mission. This configuration could be adjusted further using communication via walkie-talkies with the Control Center for clarifications. The protocol established for this configuration operates as follows: The Control Center initiates a current status request to an FR identifier through the MQTT broker. If there exists an MMPI associated with that identifier on the network, it responds with an MQTT message containing its present configuration and the currently displayed points of interest on the screen. These parameters are integrated into the Configuration Dashboard, allowing modification by the Control Center. Subsequent alterations can then be transmitted to the FR using a new MQTT message.

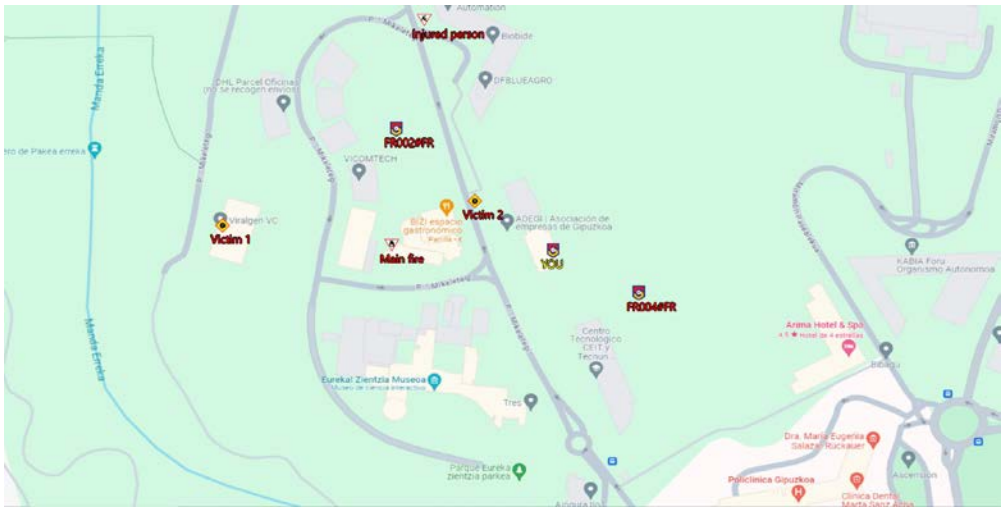


Figure 3. Map view

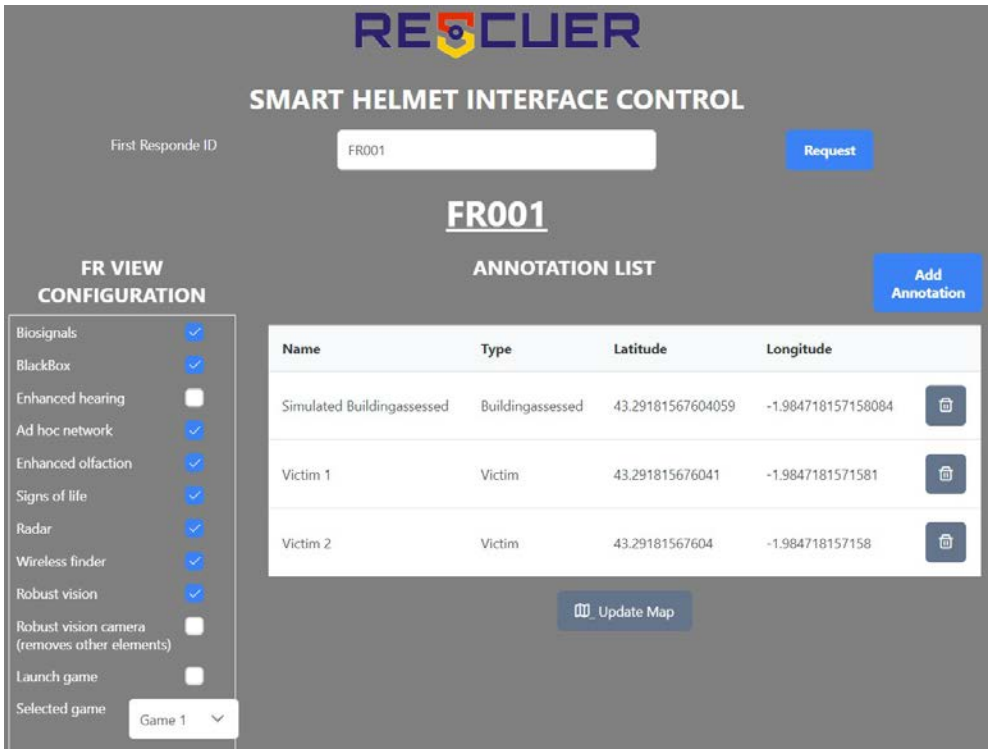


Figure 4. Interface configuration dashboard accessible from the Command Center

Smart Helmet

The Smart Helmet is a concept introduced in Ferna´ndez Garc´ıa et al., 2023 that refers to a regular FR helmet that has been enhanced with a display, a camera, and a processing unit in order to create a tool that can transmit information to the FR in the less invasive way possible. Each of these components has the following characteristics.

Head Up Display

A micro OLED transparent one-eyed display, integrated in a non-intrusive manner into the existing protective gear of the FR, capable of projecting AR information congruent with the environment. It includes a microcontroller in charge of controlling the brightness and orientation of the display. The HUD is mounted on a 5-degree-of-freedom mount that can be adjusted according to the user's ergonomics. This mount is affixed on the front side of the strap attached to the helmet.



Figure 5. Draft concept of the system installed on user's helmet and the final prototype in Ferna´ndez Garc´ıa et al., 2023

Helmet Mounted Camera

In some situations, such as search and rescue, cameras operating at different wave frequencies, such as infrared or ultrasound, are used to penetrate through debris or adverse conditions and detect the presence of trapped people. Similarly, thermal cameras detect heat emitted by objects and people rather than relying on visible light. They are essential in situations such as fires, earthquakes, or avalanches, as they can detect bodies or heat sources even in conditions of smoke, darkness, or debris. In this case, as the objective is to verify the transmission of the frames received from any type of camera, the considered prototype integrates a GoPro, and the flow of the received frames.

Processing Unit (Edge Unit) and an External Battery

Lightweight, energy-efficient, and integrated into the FR protective gear, it will provide an efficient processing platform with capabilities to develop rendered camera images and auxiliary sensory information, resulting in an AR processing unit.

Testing in Realistic Scenario

This prototype was tested as part of the RESCUER, 2023 project, where a series of pilots were planned in which the FRs tried first-hand the technology developed in the project in different use case scenarios. Three of these international events were located in Weeze (Germany) where an earthquake scenario was simulated, Navacerrada (Spain) for mountain rescue operations, and Modane (France) for tunnel rescue situations, in the consecutive months of November, January, and October of the period 2022-2023. Among other technologies, the AR interface was tested as a displaying tool for the Smart Helmet. These tests include a first phase to train the FRs to familiarize them with the interpretation of the display. In the second phase, the FRs operated in a simulated Emergency Field and performed a series of rescue tasks: dragging victims, detecting and transmitting the presence of dangerous products, and rescuing in smoke and darkness, among others. Finally, the FRs provided feedback evaluating their experience. The statistics of FRs feedback collected during the first two pilots are published in Ferna´ndez Garc´ıa et al., 2023, while the third round of Modane is in progress. This feedback has been essential to evolve the AR design according to the FRs requirements. The modules of data designed for this validation use case include the following information:

Location Information

The visualization of objectives and the positions of other First Responders (FRs) is a key parameter in rescue operations. In the RescueR project, this information is represented within an outer circle (depicted in Figure 2), illustrating the names and distances to the agents and objectives, such as other FRs, victims, and dangerous obstacles. This visual representation resembles the perspective commonly observed in video games. Notably, the orientation of the FR influences the adjustment of this circle, aligning it accordingly, and facilitating an intuitive orientation in the emergency field. For instance, when the object appears at the top of the circle, it signifies its location as being directly in front of the FR, establishing an intuitive and dynamic display mechanism for spatial awareness.

Sensor Kit Simulator

The modular development of the RescueR architecture enables simultaneous progress on various fronts, allowing each center to concentrate on advancing its own technology. Yet, particularly within a European project, coordinating for testing and validating distributed work from diverse consortium members demands considerable effort. This necessity sparked the inception of the Sensor Kit Simulator concept. This software is purpose-built to simulate messages from various sensors and tools used in the project within an MQTT broker. The paper focuses on the AR interface, where the relevance of the data simulated in the MQTT broker lies in its adherence to the original messages' structure, keywords, and frequency, rendering their origin inconsequential. To achieve this, the consortium standardized a common structure, enabling each partner to determine the requisite information to transmit from their sensors. This approach facilitated the independence of tool development and progress. The design of the icons-system as a representation of the distinct modules in Table 1 has been defined in collaboration with the *Ecole nationale supe´rieure des officiers de sapeurs-pompiers* (ENSOSP), in order to facilitate the comprehension and integration of this tool of firefighters within their rescuer tasks. Indeed, the icons in Table 1 are distinguished in two levels, since some modules provide information registered in more than one sensor. In these cases, as shown in Figure 2, the icons of the sensors are grouped in boxes, so that the symbols of the modules are in the upper left edge (Bio-signals, Black box, Ad-hoc network, and Augmented olfaction), and the symbols of the sensors (heart and breath rate, temperature, humidity, ...) inside the boxes. This distinction of these symbols is precise in Table 1: the left column contains the icons of the modules, while the right column incorporates the icons of the multi-sensor modules to the module descriptions.

In fact, the pilot of Mountain Rescue indicated the need to improve the AR interface in order to reduce the amount of information displayed and avoid distractions in the rescue tasks. To do this, one of the keys of the Rescuer project was to calculate the cognitive load and regulate the projection of information. Considering the recent results of Gutie´rrez et al., 2023, with the available sensor kit it was not possible to measure the cognitive load directly, but it is possible to derive it from bio-signals such as respiration, blood volume pulse, and electrocardiogram. Therefore, the modularity of the interface has been designed, i.e., the selected information can be projected from the entire data stream received (see Figure 6 in the Appendix for modularity explanation purposes). An automated modularity according to the cognitive load has also been considered in the project. But without robust and sufficient results to create such automation on the DSO, it was decided to use the Control Dashboard, through which, having direct communication with the FRs of the field plus their bio-signals, the Control Center operators customize the visualization of the AR. This AR has been tested in the Tunnel scenario pilot, but due to the processing status of the statistics, the present study is considered a work in progress.

CONCLUSION AND FUTURE WORK

This developed modular platform has already been tested, and the main conclusion has been that it provides real-time information to the FR that aims to facilitate their decision-making without augmenting the cognitive load of the user. Indeed, the modularity of the created interface allows the deployment of adapted information exposure respecting FRs' physiological limitations and capacities in the Emergency Field. Furthermore, the participation of the FRs in the definition of icons and colors to represent the situational awareness-enhancing modules facilitates the training process of the FRs. Integration of previously recognized symbols in rescuer vocabulary makes more innate the information understanding through an AR interface. All these characteristics made the MMPI an AR interface, in adequate circumstances, capable of augmenting the situational awareness of the FRs interfering in a minimal way in their rescuer maneuvers. However, there may also be situations in which the technology is not supportive. A blow can damage the equipment and disorient or confuse the FR, but these are inconveniences that, regardless of the technology, can not always be prevented.

One of the next objectives is the implementation of an automatized modular exposure based on the biosignals of the FRs. This smart modular adaptivity will be an optional display mode, in order to ensure that the AR does not overload the FRs and interfere with their operations. This should be an advanced Data Sharing Orchestrator version, considering the FR working team configurations in the Emergency Field and providing appropriate information distribution. For example, there may be three distinct configurations for situations such as exploration, rescue operation, and survival mode. Each one of those would provide a distinct data layout. For exploration or victim-finding operations, all FRs may receive the complete set of situational awareness modules, especially tools that facilitate the detection of people under low visibility circumstances. On the other hand, once victims are detected, for rescue operations, individual biosignals may be displayed, but there may be an information reduction in a manner that one FR may receive contextual information (temperature, humidity, and the number of surrounding people), while a team-mate obtains other specific tool derived data. In extreme situations, minimal data should be

Table 1. Symbol description of the situational awareness modules and sub-modules

Situational Awareness enhancing MODULES	Description of corresponding sensors
	Biosignals: Description of the biological signals of the FR. The color of the icons will change from green to red depending on if the values received are good or bad.  Heart rate: Measure of the heartbeat per minute.  Breath rate: Number of breaths per minutes done.
	Black Box: The black box is a device that has some environmental sensors inside; temperature, humidity and number of people around it.  Temperature: Temperature from the environment around the Black Box.  humidity: Humidity of the environment around the Black Box.  People: Number of people currently surrounding the Black Box.
	Ad-hoc network: Information of the wireless network generated by the gateways that each FR carries during an operation. It provides information about the battery level of the gateway and the status of communications between the C2 and other FRs.  Battery level: battery level of the gateway that both connects and is part of the generated network. The image is updated depending on the amount of battery left.  C2 connection: Flag showing if the control center is reachable through the Ad.hoc network. If so, the C2 could send coordinates of points of interest or modify the view on the interface.  FR connections: Number of FRs reachable within the Ad-hoc network.
	Augmented olfaction tool: It provides the concentration values on up to 5 type of gases, turning from green to yellow to red depending on if the gas levels are dangerous or not.
	Signs of life: This tools provides information if it detects life (through walls for example) it provides an estimated value of the distance to the found individual.
	Radar: This device shows the amount of objects that are approaching to the FR. The interface blinks when an object is approaching.
	Wireless finder: This device measures the distance to devices that produce wireless signals such as Bluetooth or Wi-Fi, that usually come from Mobile phones from victims under the rubble. The interface guides the FR in the initial calibration phase also.
	Robust vision: Displays the type and amount of objects detected by the object-detection module. It can be switched to a full camera view that will show the IR video feed with the red squares instead of the transparent view.

displayed in order to ensure the safety of the FRs, e.g. providing exit location points. Therefore, the priority of modules should be analyzed according to biosignals and the risky level of the circumstances of the Emergency Field.

In summary, this approach intends to replace the necessity for manual adaptation of information by the control center, thereby implementing an intelligent version of the current Modular Multi-Platform Interface.

APPENDIX

Figure 6 illustrates the reduced information exposure in the Modular Multi-Platform Interface (MMPI). This modular characteristic of the interface intends not to overload the First Responder when its cognitive load is high.



Figure 6. Example of the Modular Multi-Platform Interface (MMPI) visualization with high cognitive load and reduced information exposition

REFERENCES

Arregui, H., Irigoyen, E., Cejudo, I., Simonsen, S., Ribar, D., Kourtis, M.-A., Spyridis, Y., Stathakarou, N., & Batistatos, M. C. (2022). An augmented reality framework for first responders: The respond-a project approach. *2022 Panhellenic Conference on Electronics & Telecommunications (PACET)*, 1–6.

Camp, P. J., Hudson, J. M., Keldorph, R. B., Lewis, S., & Mynatt, E. D. (2000). Supporting communication and collaboration practices in safety-critical situations. *CHI'00 extended abstracts on Human factors in computing systems*, 249–250.

Choong, Y.-Y., & Salvendy, G. (2021). Voices of first responders–applying human factors and ergonomics knowledge to improve the usability of public safety communications technology.

Ferna´ndez Garc´ıa, A., Oregui, X., Lingos, K., Konstantoudakis, K., Belmonte Herna´ndez, A., Azpiroz, I., & Zarpalas, D. (2023, November). Smart Helmet: combining sensors, AI, Augmented Reality and personal protection to enhance first responders' situational awareness. https://doi.org/10.5281/zenodo.10141839

Frederick, S., Kahneman, D., & Mochon, D. (2010). Elaborating a simpler theory of anchoring. *Journal of Consumer Psychology*, *20*(1), 17–19.

Gutie´rrez, Á ., Blanco, P., Ruiz, V., Chatzigeorgiou, C., Oregui, X., Á lvarez, M., Navarro, S., Feidakis, M., Azpiroz, I., Izquierdo, G., Larraga-Garc´ıa, B., Kasnesis, P., Olaizola, I. G., & Á lvarez, F. (2023). Biosignals monitoring of first responders for cognitive load estimation in real-time operation. *Applied Sciences*, *13*(13). https://doi.org/10.3390/app13137368

IFAFRI. (2023). *International forum to advance first responder innovation (ifafri)*. https://www.internationalresponderforum.org/services/capability-gaps

Jain, N., Mangal, P., & Mehta, D. (2014). Angularjs: A modern mvc framework in javascript. *Journal of Global Research in Computer Science*, *5*(12), 17–23.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585.

Kapalo, K. A., Bockelman, P., & LaViola Jr, J. J. (2018). "sizing up" emerging technology for firefighting: Augmented reality for incident assessment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 1464–1468.

Ludwig, T., Reuter, C., & Pipek, V. (2013). What you see is what i need: Mobile reporting practices in emergencies. *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*, 181–206.

Meneguzzi, F., Oh, J., Chakraborty, N., Sycara, K., Mehrotra, S., Tittle, J., & Lewis, M. (2012). A cognitive architecture for emergency response.

Monteiro, P., Gonc¸alves, G., Coelho, H., Melo, M., & Bessa, M. (2021). Hands-free interaction in immersive virtual reality: A systematic review. *IEEE Transactions on Visualization and Computer Graphics*, *27*(5), 2702–2713.

Neustaedter, C., McGee, J., & Dash, P. (2019). Sharing 9-1-1 video call information between dispatchers and firefighters during everyday emergencies. *Proceedings of the 2019 on Designing Interactive Systems Conference*, 567–580.

NG911. (2023). *Next generation 911*. https://www.ng911now.org/

Park, H., Park, J., Lin, S.-H., & Boorady, L. M. (2014). Assessment of firefighters' needs for personal protective equipment. *Fashion and Textiles*, *1*(1), 1–13.

RESCUER. (2023). *First responder-centered support toolkit for operating in adverse and infrastructure-less environments*. https://rescuerproject.eu/

Long-Range Human Detection in Drone Camera Images

Joris Heemskerk
University of Applied Sciences Utrecht
joris.heemskerk@hotmail.com

Tina Mioch
University of Applied Sciences Utrecht
tina.mioch@hu.nl

Henry Maathuis
University of Applied Sciences Utrecht
henry.maathuis@hu.nl

Huib Aldewereld
University of Applied Sciences Utrecht
huib.aldewereld@hu.nl

ABSTRACT

In recent years, drones have increasingly supported First Responders (FRs) in monitoring incidents and providing additional information. However, analysing drone footage is time-intensive and cognitively demanding. In this research, we investigate the use of AI models for the detection of humans in drone footage to aid FRs in tasks such as locating victims. Detecting small-scale objects, particularly humans from high altitudes, poses a challenge for AI systems. We present first steps of introducing and evaluating a series of YOLOv8 Convolutional Neural Networks (CNNs) for human detection from drone images. The models are fine-tuned on a created drone image dataset of the Dutch Fire Services and were able to achieve a 53.1% F1-Score, identifying 439 out of 825 humans in the test dataset. These preliminary findings, validated by an incident commander, highlight the promising utility of these models. Ongoing efforts aim to further refine the models and explore additional technologies.

Keywords

Human Detection, Computer Vision, Drones.

INTRODUCTION

In the last years, the use of drones to aid First Responders (FRs) has become more and more prevalent, including applications such as monitoring fires from an unobstructed perspective (e.g., (Georgiades et al., 2019)) and assessing building damage in difficult-to-access areas (e.g., (Adams et al., 2014)). The camera footage recorded by drones can also be used, for example, to help detect victims, solve logistical problems, and locate missing persons (e.g., (Khan & Neustaedter, 2019)). Currently, FRs usually fly the drones manually and interpret the camera footage without any system help. However, analysing drone images is time-intensive and cognitively demanding, making the interpretation of the camera footage prone to errors. Automating aspects of this analysis, e.g., the detection of humans, can help support FRs in their tasks and decision-making and possibly lead to an overall better understanding of the environment and reduce the time needed to provide aid to victims.

Past research focused on detecting objects like roofs, cars, debris, and vegetation from cameras in UAVs or helicopters (Pi et al., 2020) and also on detecting missing people in for example nature environments (Martinez-Alpiste et al., 2021). We aim to expand upon this by introducing more complex scenarios, like residential and industrial areas, with drones flying at varying heights, angles, and scales. For this, we created a custom dataset based on drone footage from the Dutch Safety Region *Rotterdam-Rijnmond*. This paper explores the use of a state-of-the-art object detection model for real-time long-range human detection in crisis scenarios. It should be noted that this explicitly differs from human identification, in which the goal is to exactly recognize individuals, and thus explicitly excludes surveillance. We only aim to locate humans present in the camera footage, and only during incidents. As such, the central question guiding this research is: *How can humans be located in drone camera footage in order to positively impact the decision-making of first responders?* As we are still in progress of investigating this question, our paper presents first steps and preliminary results.

WiP Paper – IT Solutions for Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Identifying humans in drone data can benefit the decision-making process of FRs in different ways. First of all, locating and marking all humans present can help to determine whether humans require rescue or are located in potentially hazardous areas. Secondly, knowing the total number of humans present at a given moment can assist the FRs in addressing logistical challenges, such as determining the number of humans requiring evacuation or shelter. We analyse the preliminary findings together with an incident commander of the *Safety Region Rotterdam-Rijnmond*, to analyse the impact the current leading model can have on their decision-making process.

In this paper, we first explore some of the related background information and describe how we created and prepared our dataset, our model choice, and our evaluation methods. After this, we present our preliminary results, both the quantitative performance as well as qualitative results of an evaluation by an incident commander from the *Safety Region Rotterdam-Rijnmond*. Finally, we discuss our findings and describe future work.

BACKGROUND

In recent years, advancements in Computer Vision and Machine Learning have revolutionized the field of object detection (Zaidi et al., 2022), contributing to innovative solutions in various domains. Specifically, automatically detecting humans in images has been popular (Nguyen et al., 2016). Most of the time, this regards detecting humans from a short-range perspective at a relatively close distance (e.g., from CCTV footage or handheld mobile phone pictures). This is reflected by widely used datasets such as COCO (Lin et al., 2014), which contains many images of, amongst others, humans, from different angles, backgrounds, and poses. However, in these datasets, humans always make up a large part of the image by size. For our research, this is not the case, as humans in our footage take up a very small part of the images, meaning there are fewer pixels to describe the humans with, making their detection more difficult. There are also various long-range datasets. However, most of those do not contain humans (e.g., MOHR (Zhang et al., 2021) and DroneVehicle (Sun et al., 2022)). An example of a dataset that does allow long-range human detection is TinyPerson (Yu et al., 2020), which consists of helicopter footage, recorded above crowded beach scenes. Whilst the scale of the humans in the drone camera footage recorded by FRs resembles this dataset, the humans in the TinyPerson dataset are sometimes even smaller than the ones in our drone footage. In addition, the beach backgrounds in the TinyPerson dataset do not resemble ours, and they vary relatively little compared to the footage recorded by FRs.

Although these datasets have different characteristics compared to our dataset, we can still learn from these applications to inform our model choices. For example, Tang et al. (2023) trained a Faster R-CNN model along with several YOLO models on the TinyPerson dataset, achieving their best Average Precision of 9.5% with their modified version of YOLOv7. Additionally, YOLOv5 and Faster R–CNN have been trained and compared on drone footage recorded in nature backgrounds (Bachir & Memon, 2024), achieving mAP 0.5 scores of 96.9% and 91.0% respectively. Lastly, a skip-based Fully Convolutional Network (FCN) was trained on a set of long-range infrared images, resulting in an F1-Score of 98.93% (Haider et al., 2021). For this research, we analyse the performance of the YOLOv8 model, a leading model in the YOLO series.

DATA

In this paper, we train and evaluate our models on a custom-made dataset based on the drone footage of the *Safety Region Rotterdam-Rijnmond*. They have a historical record of some of the incidents where drones were used, going back to 2019. The footage is recorded in Full HD (1920 by 1080 pixels), 30 frames per second, by Zenmuse H20N cameras[1] on DJI Matrice 300 RTK drones[2]. The data contains a lot of variation, e.g., regarding light conditions (day and night), environmental background (e.g., industrial areas, rural, urban, indoor), and camera angles and distance (see Tables 1 and 2). For the creation of our dataset, we reduced the variation in the backgrounds by selecting exclusively daytime footage in industrial and residential areas, keeping only the video footage that was not zoomed in. From this, we only kept the 12 scenarios that contained humans and did not exclusively contain FRs. In the Methodology Section, we elaborate on the methods we used for preparing and labelling the footage from these scenarios to create a dataset. For examples of the selected footage, see Figure 1. In later stages of this research, we aim to incorporate more varied scenarios or train separate models for different situations.

METHODOLOGY

In this section, we go over the different steps we took to get to our preliminary results. Firstly, we elaborate on the methods used to prepare our custom dataset. Secondly, we explain our labelling process, after which we introduce our model choice. Lastly, we describe our evaluation set-up, where we differentiate between a quantitative and a qualitative evaluation.

[1] https://enterprise.dji.com/zenmuse-h20n
[2] https://enterprise.dji.com/matrice-300

Table 1. Scenario frequency for all available data. Some incidents span multiple categories, in which case they are counted in all.

Scenario category	Frequency
Industrial	67
Nature	50
Water	48
Urban	41
Indoor	8

Table 2. Light condition frequency for all available data.

Light condition	Frequency
day	104
night	53



Figure 1. Example images from the selected incidents, that cover a range of camera angles, backgrounds, and scales.

Data Preparation

In this section we explain the steps we took to go from the footage of the 12 selected incidents, as discussed in the Data Section, to an image dataset, containing 5,357 images. We also elaborate on the methods used to reduce this dataset into various representative and more balanced subsets.

The 12 incidents combined, contain a total of 87 videos with an average runtime of 7.5 minutes. We converted these videos into a set of images. We extracted the images at an interval of 150 frames (equivalent to 5 seconds) to mitigate redundancy, because the footage does not differ much from frame to frame, as the drones often move slowly or are stationary. This results in a total of 6,694 pictures. By filtering out the images without humans, 5,357 images were left (80.0%). We still noticed an overlap in some of the images, as well as an imbalance in the remaining data, as some incidents are over-represented by the number of images (due to different lengths of the original video footage).

In order to solve the imbalance problem, whilst simultaneously mitigating the perceived image overlap, we balanced the dataset by using a technique that reduces the entire dataset into a subset of x images. This needs to be done in such a way that the resulting subset is representative of the original dataset. A way to do this would be to not group the images by the scenario they stem from to create the reduced subset, but to cluster the data by using the backbone of a classification model, i.e., using a classification model to embed each image into a feature space, where all images that are similar as seen by the model will be close to each other. A classification model extracts certain features from an image, e.g., the relative frequency of green lines or the redness of the image, such that we can group the images on these features. A frequently used classification model is ResNet-50 (He et al., 2016). This model is pre-trained on the ImageNet 2012 classification dataset (Russakovsky et al., 2015), a dataset containing 1.28 million training images spread across 1000 classes, including animals, vehicles, common household objects, and more. The ResNet-50 backbone embeds each image into a 2,048 feature space. In this space, images that are close to each other are considered to be similar by ResNet-50. The classification step that follows the embedding is not needed, as the label does not help reduce the dataset.

Figure 2 is a visualisation of how all images with people were clustered, using t-distributed Stochastic Neighbour Embedding (t-SNE) (Van der Maaten & Hinton, 2008), which is a technique that reduces larger feature spaces into just two dimensions, trying to keep data points close together that are also close in the original space. It is clearly

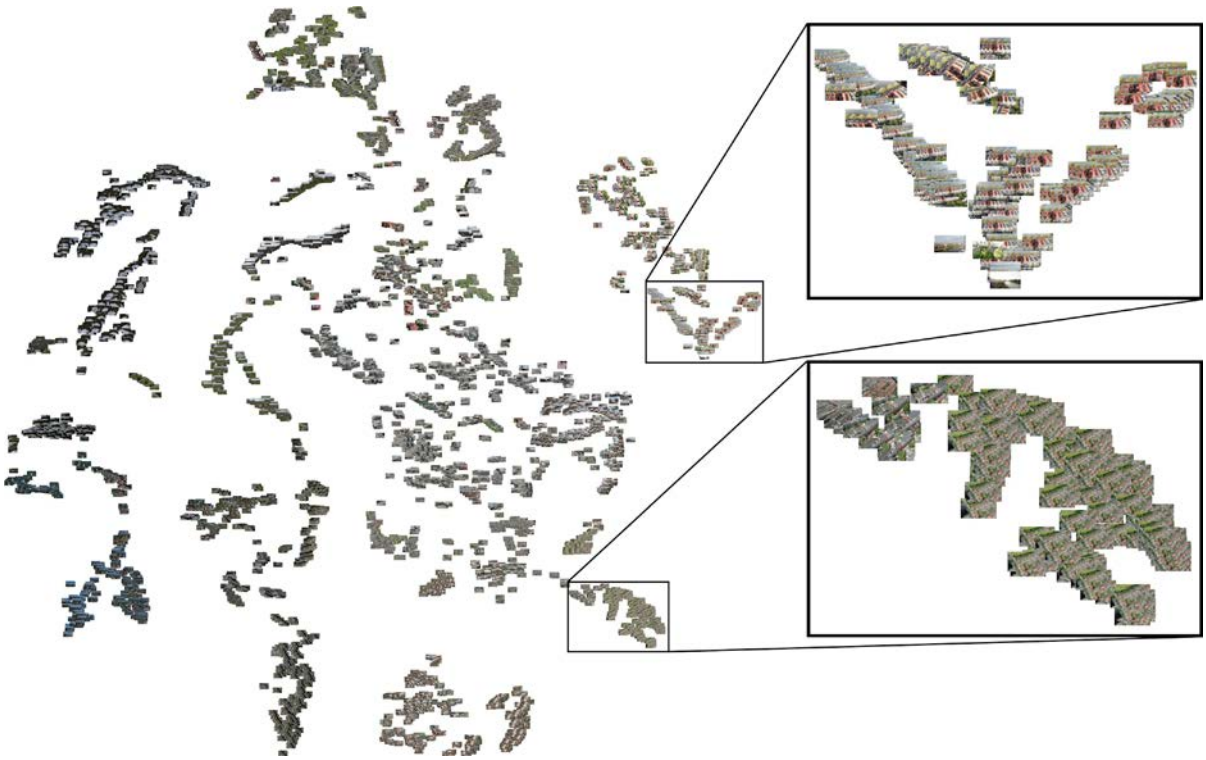


Figure 2. A t-SNE representation of the feature space, made by ResNet-50 on all the fire brigade's images that contain people.

visible in the zoomed-in regions of Figure 2 that the images with very similar structured content are close together and that the two regions are distinct from each other.

To generate a representative subset, the following heuristic was used. First, a random point was selected in the original feature space, after which the furthest possible point was selected as the second point. For this, the cosine similarity was used, to prevent the subset from primarily consisting of outliers. The third point is defined as the point that is the furthest possible point from both points, again using the cosine similarity. The next point is the furthest from all former points, and so on. Using this heuristic, multiple formations are repeatedly created, each with a different starting point. The formation that has the highest minimum cosine similarity will be kept. This is the formation that consists of the points that are the furthest possible from each other.

Using this technique, we created multiple subsets, decreasing in size, such that each subset is created from the closest bigger subset. This will make us able to easily increase the number of images in later stages of this research, knowing that the images we already prepared and labelled can be reused in bigger subsets. In Figure 3, a subset of 100 images, as selected from a subset of 200 images, is visualised in the entire feature space.

Data Labelling

For this paper, we have labelled a subset of 100 images. An example of a labelled image can be found in Figure 4. In the case of a crowd, we attempted to label every single human individually. We also labelled partially obstructed people, as they appear frequently and their detection might prove just as vital when it comes to detecting for example victims.

Model Choice

We trained the Machine Learning model YOLOv8[3] to detect humans in the fire brigade's footage. YOLOv8 is being regarded as state-of-the-art (Solawetz & Franchesco, 2023). Due to its multi-scale prediction method and anchor-free architecture (Jocher et al., 2024), it is specifically good at detecting small-scale objects, making it very applicable to the problem at hand. YOLOv8 is pre-trained on the COCO dataset (Lin et al., 2014). In this paper, we will fine-tune the pre-trained model.

[3] https://github.com/ultralytics/ultralytics

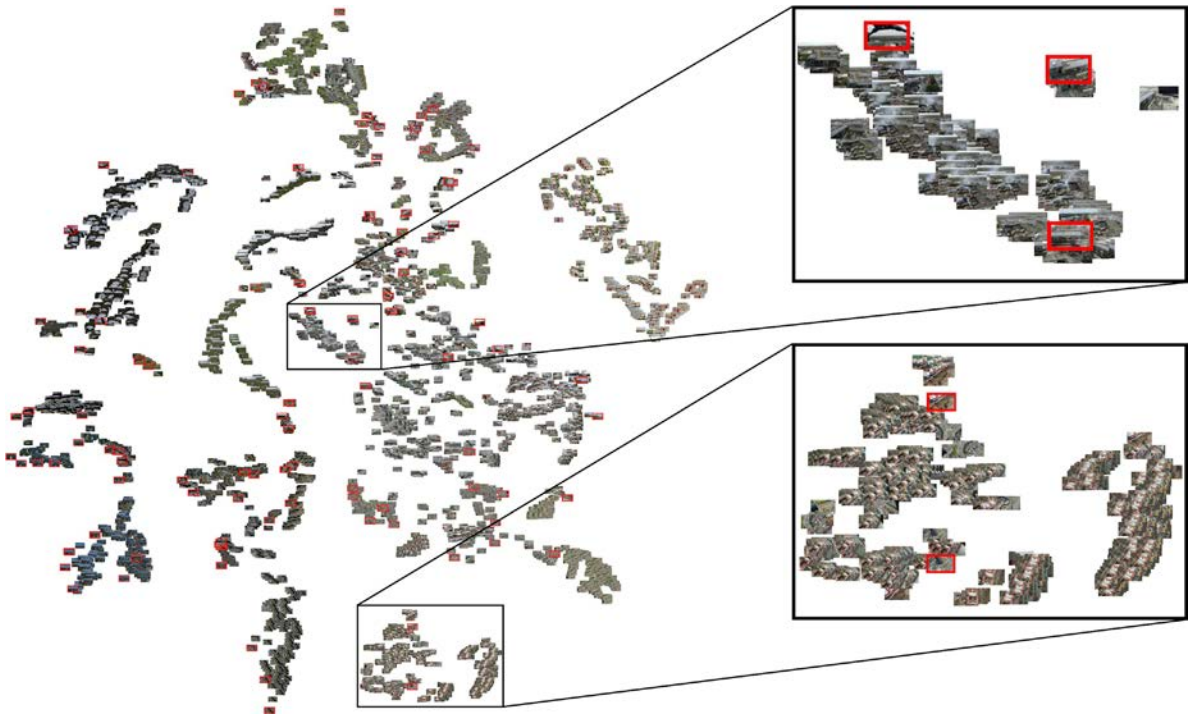


Figure 3. A t-SNE representation of the feature space, where the images with the red borders denote them as selected in the subset of 100.



Figure 4. An image with the labels overlaid. We zoomed in on some of the labelled humans.

Various versions of YOLOv8 are available, including *nano*, *small*, *medium*, *large*, and *extra large*, referring to the size of the model and its parameters. The larger the model, the more information it can store. A larger model has the capacity to learn more detailed representations of what a human is, but could also more easily overfit on smaller datasets. For our research, we decided to train these different versions and compared their results. We also explored the effects of different batch sizes, i.e., the number of images the model is exposed to before adjusting its weights. Larger batch sizes reduce the likelihood of the model overfitting the training data, but can negatively affect how long it takes the model to converge. In addition, we varied the input resolutions of the model, which have to be a multiple of the default 640 by 640 pixels, meaning we were only able to additionally analyse 1280 by 1280 pixels. The other parameters were left as default. The models' initial learning rates were 0.01. The stochastic gradient descent (SGD) optimization algorithm was used with a default momentum of 0.9 and attenuation coefficient of 0.0005. A constant seed of 42 was set for reproducibility purposes. The default mosaic augmentations were used. Each model was trained for 100 epochs, as we found this to be enough for all of the models to converge. The best scoring model, based on F1-Score, out of all these epochs was kept. Lastly, we randomly divided the 100 labelled images into a training, test, and validation set with a ratio of 8:1:1.

Evaluation Set-up

To evaluate our models, we use both quantitative and qualitative evaluation methods, which we describe in the following sections.

Quantitative Evaluation

To evaluate the performance of the model, we considered Precision, Recall, F1-Score, and mAP (Mean Average Precision). The precision and Recall of the model are calculated as follows:

$$Precision = \frac{TP}{TP + FP}, \qquad Recall = \frac{TP}{TP + FN},$$

where TP = True Positive, FP = False Positive, and FN = False Negative. Precision here quantifies the ratio of correctly predicted humans. Recall quantifies the proportion of actual humans that were correctly identified.

We combine recall and precision into an F1-Score to get a single score providing insight into both how often the model correctly identifies humans and how many humans it misses. A high F1-Score means the model is both accurate (with minimal False Positives) and thorough (with minimal False Negatives) in finding humans in images:

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}.$$

Since this score provides a balanced measurement of both False Positives and False Negatives, we primarily used this metric to intercompare models. For the best-performing model, based on the F1-Score, we will also analyse the exact number of True Positives, False Positives, and False Negatives, with the use of a confusion matrix. This gives us a more accurate and exact number to analyse if the best-performing model is business-acceptable or not.

Lastly, we used the mean Average Precision metric (mAP), which is commonly used in assessing the performance of object detection models. The Average Precision (AP) summarises the shape of the precision/recall curve, which shows how precision and recall change as we vary the threshold for what counts as a "correct" detection. The area under this curve gives us the average precision. This is done by taking the precision at each recall point and then averaging over all recall values. Mathematically, it's represented as:

$$AP = \sum_{r \in \text{recalls}} (r_i - r_{i-1}) \cdot p_i,$$

where r_i and r_{i-1} are consecutive recall values and p_i is the precision corresponding to the recall r_i.

The mAP is then computed as the average of AP over all classes:

$$mAP = \frac{1}{N} \sum_{k=1}^{N} AP_k,$$

where N is the number of classes and AP_k is the AP of class k. The mAP is a measure of accuracy for object detection models, typically assessed at various Intersection over Union (IoU) thresholds. IoU represents the percentage of overlap between the predicted bounding box and the ground truth bounding box:

$$IoU = \frac{A \cap B}{A \cup B},$$

where A is the actual bounding box and B is the bounding box as predicted by the model. For example, a mAP 0.5 score indicates the Mean Average Precision at an IoU threshold of 50%.

Qualitative Evaluation

Whilst these quantitative evaluation metrics enable us to intercompare models, they do not allow us to evaluate if any single model is business acceptable, i.e., can positively impact the decision-making of FRs. Relevant questions concern, for example, the potential effects of people being overlooked in the footage (False Negatives), people being marked multiple times, background objects being marked as people (False Positives), and the confidence values given for each detection. As a first step towards evaluating the model, we discussed the model results with an incident commander by presenting different scenarios with corresponding model output. The scenarios we presented differed in the number of humans that were present (e.g., a few humans or a crowd) and the setting of the incident (e.g., an industrial or residential area). The participant was asked to discuss the model output, including the impact on decisions, the desirability of different model behaviour, and other aspects relating to the understandability and usability of the model.

PRELIMINARY RESULTS

In this section, we describe the quantitative and qualitative preliminary results.

Quantitative Results

In this section, we present the preliminary results of the experiments we have run so far. We analyse the accuracy of all the different YOLOv8 models (i.e., *nano*, *small*, *medium*, *large*, and *extra large*) for a varying number of batch sizes and input resolutions. The training and testing was performed on a HP ZBook Studio 15.6 G8 laptop[4] with a Nvidia RTX A3000 GPU with 6GB of VRAM (Video Random-Access Memory). The relatively limited VRAM resulted in certain tests not being runnable, which explains why not every model is trained on every batch size.

640 by 640 resolution results

In order to analyse the effect different batch sizes have on the performance of the different models, we first ran them all with a 640 by 640 pixel input resolution, as we were not able to run larger batch sizes with higher input resolutions. This means that the images are scaled down, losing some detail in the progress. In return, we were able to train more model configurations. Due to potentially promising results, the YOLOV8s model on a batch size of 16 was trained on CPU.

In Table 3 we see that, judging by the F1-Scores, the two best-scoring models are YOLOv8s with a batch size of 12 and YOLOv8m with a batch size of 8. We do see a difference in Precision and Recall between the models. The YOLOv8s model seems to have a lower Precision score, but a higher Recall, which means that this model has more True Positives, at the cost of some extra False Positives.

1280 by 1280 resolution results

Due to our expectation that a higher input resolution would lead to better results, as there would be more pixels to describe the people with, we also tested the 1280 by 1280 input resolution. Not many of the tests would run on our hardware, which is why the larger batch sizes for the YOLOv8n model were trained on the CPU. In order to have a result for each model, we also introduced a batch size of 1. The 1280 by 1280 input resolution results in an overall higher performance compared to the 640 by 640 input resolution results (see Table 4). The YOLOv8n models show higher scores than the other models, with the batch size 16 model scoring the best in its F1-Score.

The validation set consists of 825 labelled people. This means that this is the worst possible False Negative score that can be reached, as well as the best possible True Positive score. In Table 5 we see that more than half (439/825) of the people are found. However, the model also marks 276 False Positives.

Qualitative Results

Based on the experimental results, for this evaluation, we used the output of the YOLOv8n model that was trained with an input resolution of 1280 by 1280 with a batch size of 16. During the discussions of the different scenarios and model outputs, the participant mentioned that it is important to differentiate between two aspects. First, the model can create False Positives, i.e., mark humans in the images where there are no humans (see examples in Figure 5). The participant mentioned that in general, False Positives should not be problematic, as having more False Positives probably also corresponds to having more True Positives and because it is better to find too many humans than too few, given that finding humans corresponds to the overall goal of the system. Also, the FR can

[4] https://support.hp.com/us-en/document/ish_4225460-4216961-16

Table 3. Comparative results of the varied YOLOv8 models with different batch sizes for an input resolution of 640 by 640 (%). The best-performing models, based on F1-Score, are highlighted in bold.

Model configuration		Performance metrics (%)			
Model name	batch size	Precision	Recall	F1	mAP 0.5
YOLOv8n	2	19.8	15.5	17.4	8.4
	4	24.3	16.6	19.7	10.1
	8	29.5	15.5	20.3	11.2
	12	25.5	20.2	22.6	11.6
	16	27.4	16.2	20.4	10.3
YOLOv8s	2	7.5	17.8	10.6	8.3
	4	24.1	17.0	19.9	10.9
	8	25.9	22.7	24.2	14.3
	12	**26.8**	**23.0**	**24.8**	**14.7**
	16	25.6	22.5	24.0	14.5
YOLOv8m	2	5.8	13.3	8.1	5.8
	4	37.0	18.4	24.6	16.3
	8	**34.6**	**19.3**	**24.8**	**16.4**
YOLOv8l	2	0	0	0	0
	4	7.9	13.1	9.8	6.6
YOLOv8x	2	7.1	3.6	4.8	1.9
	4	6.2	13.9	8.6	5.8

easily (visually) verify the result in the image. Second, the model can create False Negatives, i.e., not detecting humans in the image (see examples in Figure 6). The participant mentioned that this is more problematic, though it might not have a big impact if the humans are clearly visible. In any way, the participant mentioned that the model already provides output that is very helpful and would like to use it.

DISCUSSION

In this paper, we explored the use of state-of-the-art Object Detection models to detect people in long-range drone footage, to aid the decision-making process of FRs. We looked at the performance of different versions of the YOLOv8 model, for different batch sizes and input resolutions. Our best model (YOLOv8n with a batch size of 16 and an input resolution of 1280 by 1280 pixels) was able to detect 439/825 people in the test data, with 276 False Positives. This model had an F1-Score of 53.1%. An incident commander of the Dutch Fire Services feels confident, based on these preliminary results, that this model can aid the FRs' decision-making process for different use cases, such as locating people in the incident area faster and with less cognitive effort and aiding with decision-making regarding the deployment of teams and logistical efforts. The model can only be used to locate humans in the camera footage and cannot be used to identify (and monitor) individuals; furthermore, the model will only be used during incidents.

As a first start in this research, we reduced the variation in the drone footage and selected residential and industrial footage in daytime. In future iterations, we would like to expand the current model to include more different types of scenarios regarding environmental background and other light conditions. If this proves challenging, we alternatively will explore making specialized models for specific use cases or adding other sensor data, such as infrared.

Also, for now, we decided not to look at the effect of introducing images with no humans into the dataset. This may have caused the large number of False Positives of our model, as the model may have learned the number of people that should be present in each image. In the next steps, we will introduce different amounts of images without humans present to analyse the effect on the number of False Positives.

To create representative subsets of the selected dataset, we used a heuristic. Whilst we believe our approach to be grounded and methodical, we recognise that this approach has not yet been validated. It may also be the case that a different approach would result in a more applicable model. In the future, we will validate our approach and compare it to different techniques, using an already labelled dataset, to back up our claims.



Figure 5. Example predictions of the model. Each red box is a detected human, the number stands for the prediction certainty. We zoomed in on some False Positives (rooftop objects recognised as humans).



Figure 6. Example predictions of the model. Each red box is a detected human, the number stands for the prediction certainty. We zoomed in on some False Negatives (unfound humans by the model).

Table 4. Comparative results of the varied YOLOv8 models with different batch sizes for an input resolution of 1280 by 1280 (%). The best performing model, based on F1-Score, is highlighted in bold.

Model configuration		Performance metrics (%)			
Model name	batch size	Precision	Recall	F1	mAP 0.5
YOLOv8n	1	38.8	31.0	33.2	21.2
	2	42.6	38.3	40.3	27.5
	4	49.7	44.0	46.7	38.7
	8	54.1	47.8	50.7	43.3
	12	55.8	48.6	51.9	43.7
	16	**58.8**	**48.5**	**53.1**	**43.8**
YOLOv8s	1	16.2	34.8	22.1	20.7
	2	13.9	38.1	20.3	21.8
YOLOv8m	1	37.6	38.7	38.1	26.8
	2	42.8	23.1	30.0	22.7
YOLOv8l	1	15.6	35.4	21.7	16.9
YOLOv8x	1	7.5	25.1	11.5	10.7

Table 5. Confusion matrix of the YOLOv8n model with a batch size of 16 and an input resolution of 1280 by 1280 pixels on the test dataset containing 825 labeled humans.

		Predicted Classes	
		Human	Background
Actual Classes	Human	439	386
	Background	276	-

In addition to selecting the data, we also labelled it. In the current stage of our research, we were only able to label 100 images. We recognise that this is not a lot. We expect the addition of extra labelled data to have a significantly positive impact on the accuracy of the models. We will continue to label data until we see diminishing returns and also explore the use of active learning to aid our labelling process. With this technique, an AI model decides which images should be next to be labelled. Another technique that we think could aid us in our labelling process is the use of semi-automated labelling. An Object Detection model like the one we built could provide the first iteration of labels, after which we would only need to correct the output. This could save a large amount of time spent on labelling images.

As mentioned above, we only had access to relatively limited VRAM to run our models and for that reason could not train every model on every batch size. Current efforts aim to access better hardware so that all of the models can be run to compare performances.

Moreover, for this research, we only analysed YOLOv8 models; we realize that other models might also perform well. In the future, we will explore different models, such as R-CNNs and Vision Transformers, and see whether these provide better alternatives.

Lastly, for this research, we were only able to speak with one incident commander, which means that our results might not accurately reflect the actual implications of our preliminary results. As next steps, we aim to expand upon this evaluation by interviewing multiple stakeholders from a variety of FR functions and by evaluating the use of the model in the field.

ACKNOWLEDGMENTS

We would like to thank the *Safety Region Rotterdam-Rijnmond* for their participation in this research. This research was supported by the University of Applied Sciences Utrecht (HU) through a '*promotievoucher*'.

REFERENCES

Adams, S. M., Levitan, M. L., & Friedland, C. J. (2014). High resolution imagery collection for post-disaster studies utilizing unmanned aircraft systems (uas). *Photogrammetric Engineering & Remote Sensing*, *80*(12), 1161–1168. https://doi.org/10.14358/PERS.80.12.1161

Bachir, N., & Memon, Q. A. (2024). Benchmarking yolov5 models for improved human detection in search and rescue missions. *Journal of Electronic Science and Technology*, 100243. https://doi.org/https://doi.org/10.1016/j.jnlest.2024.100243

Georgiades, G., Papageorgiou, X. S., & Loizou, S. G. (2019). Integrated forest monitoring system for early fire detection and assessment. *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, 1817–1822. https://doi.org/10.1109/CoDIT.2019.8820548

Haider, A., Shaukat, F., & Mir, J. (2021). Human detection in aerial thermal imaging using a fully convolutional regression network. *Infrared Physics & Technology*, *116*, 103796. https://doi.org/https://doi.org/10.1016/j.infrared.2021.103796

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 770–778.

Jocher, G., Laughing-q, Chaurasia, A., & Cagatay Akyon, F. (2024). *Yolov8*. Retrieved February 6, 2024, from https://docs.ultralytics.com/models/yolov8/

Khan, M. N. H., & Neustaedter, C. (2019). An exploratory study of the use of drones for assisting firefighters during emergency situations. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3290605.3300502

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dolla´r, P., & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

Martinez-Alpiste, I., Golcarenarenji, G., Wang, Q., & Alcaraz-Calero, J. M. (2021). Search and rescue operation using uavs: A case study. *Expert Systems with Applications*, *178*, 114937. https://doi.org/https://doi.org/10.1016/j.eswa.2021.114937

Nguyen, D. T., Li, W., & Ogunbona, P. O. (2016). Human detection from images and videos: A survey. *Pattern Recognition*, *51*, 148–175. https://doi.org/https://doi.org/10.1016/j.patcog.2015.08.027

Pi, Y., Nath, N. D., & Behzadan, A. H. (2020). Convolutional neural networks for object detection in aerial imagery for disaster response and recovery. *Advanced Engineering Informatics*, *43*, 101009. https://doi.org/https://doi.org/10.1016/j.aei.2019.101009

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, *115*, 211–252. https://doi.org/10.1007/s11263-015-0816-y

Solawetz, J., & Franchesco. (2023). *What is yolov8? the ultimate guide*. Retrieved February 11, 2024, from https://blog.roboflow.com/whats-new-in-yolov8/

Sun, Y., Cao, B., Zhu, P., & Hu, Q. (2022). Drone-based RGB-infrared cross-modality vehicle detection via uncertainty-aware learning. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(10), 6700–6713. https://doi.org/10.1109/TCSVT.2022.3168279

Tang, F., Yang, F., & Tian, X. (2023). Long-distance person detection based on YOLOv7. *Electronics*, *12*(6). https://doi.org/10.3390/electronics12061502

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, *9*(11).

Yu, X., Gong, Y., Jiang, N., Ye, Q., & Han, Z. (2020). Scale match for tiny person detection. *Proceedings of the IEEE/CVF winter conference on applications of computer vision (WACV)*, 1257–1265.

Zaidi, S. S. A., Ansari, M. S., Aslam, A., Kanwal, N., Asghar, M., & Lee, B. (2022). A survey of modern deep learning based object detection models. *Digital Signal Processing*, *126*, 103514. https://doi.org/https://doi.org/10.1016/j.dsp.2022.103514

Zhang, H., Sun, M., Li, Q., Liu, L., Liu, M., & Ji, Y. (2021). An empirical study of multi-scale object detection in high resolution uav images. *Neurocomputing*, *421*, 173–182. https://doi.org/https://doi.org/10.1016/j.neucom.2020.08.074

From Data to Action: A Graph-Based Approach for Decision Support in Civil Protection Operations Planning

Sabine Janzen*
German Research Center for Artificial Intelligence, Saarbrücken, Germany
Sabine.Janzen@dfki.de

Natalie Gdanitz
German Research Center for Artificial Intelligence, Saarbrücken, Germany
Natalie.Gdanitz@dfki.de

Merlit Kirchhöfer
Federal Agency for Technical Relief, Bonn, Germany
Merlit.Kirchhoefer@thw.de

Tobias Spanke
Federal Agency for Technical Relief, Bonn, Germany
Tobias.Spanke@thw.de

Wolfgang Maaß
German Research Center for Artificial Intelligence; Saarland University, Saarbrücken, Germany
Wolfgang.Maass@iss.uni-saarland.de

ABSTRACT

In the face of increasing frequency and severity of crises such as natural disasters, pandemics, and geopolitical conflicts, civil protection organizations are crucial for recovery and support for affected populations. However, the efficiency and effectiveness of these organizations in crisis response are often hindered by the manual generation of operation plans, characterized by cognitive overload and limited analytic overview. Existing systems focus on detecting crises or post-crisis analysis, overlooking proactive planning. We present GRETA, a graph-based operational planning approach utilizing semantic historical data for better decision support. GRETA uses Operational Scenario Patterns to model operations, mapping them onto a knowledge graph in JSON-LD format, thus creating a structured representation of past data to improve future crisis response planning. We tested GRETA with Germany's Federal Agency for Technical Relief, analyzing over 157,450 historic operations from 2012-2022. Results show GRETA enhances plan efficiency, accuracy, and comprehensiveness, aiding especially inexperienced planners.

Keywords

Operations Planning, Civil Protection, Decision Support Systems, Knowledge Graph, Operational Scenario Patterns

INTRODUCTION

In recent years crises like natural disasters, pandemics or political conflicts are becoming more frequent with critical impacts on people's private and public lives, finances, economy and politics, for instance the COVID-19 pandemic, the Turkey–Syria earthquakes in 2023 or the Russian Ukrainian war. In such crisis situations, civil protection organizations like the German Federal Agency for Technical Relief (THW), French Civil Protection, Corps of the Italian Civil Protection or the Austrian Red Cross restore order, rebuild affected areas and provide support for civilians and crisis victims (Chehade et al., 2018; Habib et al., 2018; Idris and Soh, 2014; Wrona et al., 2019). To ensure efficient and effective operations on-site, accurate and complete operation plans for task-driven allocation of

*corresponding author

resources and qualified personnel are needed as fast as possible (Idris and Soh, 2014). Currently, those operation plans are generated manually by experts including all related disadvantages of human cognitive overload under time pressure as well as limited analytic overview due to heterogeneous large-scale data (Correia et al., 2018). Up til now, data-driven decision support systems are not applied within operations planning of civil protection organizations in Germany.
This leads to an inefficient use of resources, delayed responses by civil protection organizations and, in the worst case, less saved lives (Ali, 2015; Idris and Soh, 2014).
Existing approaches on decision support systems for operations planning focus on early detection of crisis signals (e.g, Bompotas et al., 2022), post-crisis planning recommendations (e.g., Krstikj et al., 2021; Lindner et al., 2018; Rauchecker and Schryen, 2018), logistical topics such as optimized route planning for resource deployment (Lim et al., 2017), evacuation plans (Lee et al., 2017), or usage of robots and drones (Habib et al., 2018). Supporting the upfront planning process of operations is beyond the scope of these approaches. In our research, we investigate decision support systems for operations planning for upcoming crisis events.
In this work-in-progress paper, we present GRETA - a graph-based approach for operations planning based on semantically enhanced historical data. GRETA introduces Operational Scenario Patterns that extend existing approaches on representing crisis events (e.g., Janzen et al., 2023) with respect to domain knowledge about operations in civil protection. Those patterns are used for mapping and aggregating large-scale data on historical crisis operations onto a knowledge graph operationalized on JSON-LD. Framed as a kind of episodic knowledge representation of past operations, the resulting knowledge graph can be used in a decision support system to generate recommendations for operation plans for upcoming crisis events.
GRETA was instantiated and evaluated in a concrete use case with the civil protection organization THW addressing planning officers on regional and supraregional levels. The applied data set covered #157,450 historic THW operations within Germany (2012-2022), including details on crisis events with respect to location and type, requester and planner of operation, date, planned duration, report on tasks, number of helpers etc. Within the case study, we showed the potential of GRETA for supporting operations planning in civil protection in terms of required time (efficiency), accuracy and completeness of resulting plans (effectiveness) as well as the empowerment of less-experienced planning officers.

RELATED WORK

Related work on graph-based decision support in crisis management (Kontopoulos et al., 2018) focus on semantic representations of specific crisis domains, e.g., fire, hazard crises or earthquakes, but not on operations (Chehade et al., 2018; Gaur et al., 2019; Hassan and Chen-Burger, 2016). Existing ontologies in the crises domain like EMERGEL[1] and MOAC[2] are not available online. Other graph-based approaches focus on natural language processing for knowledge extraction from unstructured operation descriptions in the context of military operations Zhao et al., 2022. The generation of recommendations for upcoming operations with respect to required personnel, equipment, tasks, duration etc. is beyond the scope of those approaches; as well as the usage of structured and unstructured data in decision-making for improving operations planning and thereby resilience of civil protection organizations (Shittu et al., 2018).

GRAPH-BASED APPROACH FOR OPERATIONS PLANNING IN CIVIL PROTECTION

We present GRETA, a graph-based approach for operations planning in civil protection consisting of four main modules (cf. Fig. 1): **Mapper, Neo4j Inserter, Query Builder,** and **Recommender**. Core concept of GRETA are graph-based operational scenario patterns following Chehade et al., 2018; Gaur et al., 2019; Hassan and Chen-Burger, 2016; Janzen et al., 2023; Kontopoulos et al., 2018; Mescherin et al., 2013 (cf. Fig. 2). Fig. 3 shows all entities and attributes of the operational scenario pattern as well as applied concepts of foundational ontologies and the transfer to domain-specific items in the use case (THW data). As an **identifier** of an operation, an ID or the date of the operation is used. In addition to a description of the crisis event, the pattern covers a **context** entity for specifying the event type of the crisis (e.g., natural distasters) and the concrete event (e.g., flood) (cf. Fig. 2). As defined in Janzen et al., 2023, the operational scenario pattern includes a **provenance** entity for capturing the source of crisis information and operation request as well as an entity for the **location** of the operation. If an operation was part of a greater project, this information is represented within the precondition attribute in the **reason** entity (cf. Fig. 3). **Impact** of an operation is described by the attributes duration of the operation, post-conditions (e.g., final report of operation) and the severity of the operation. The **actors** entity covers information on the total amount of actors (helpers) of the operation, their conducted working hours as well as deployed units with their strengths, roles

[1] http://disaster-fp7.eu/sites/default/files/D3.22.pdf
[2] http://observedchange.com/moac/ns

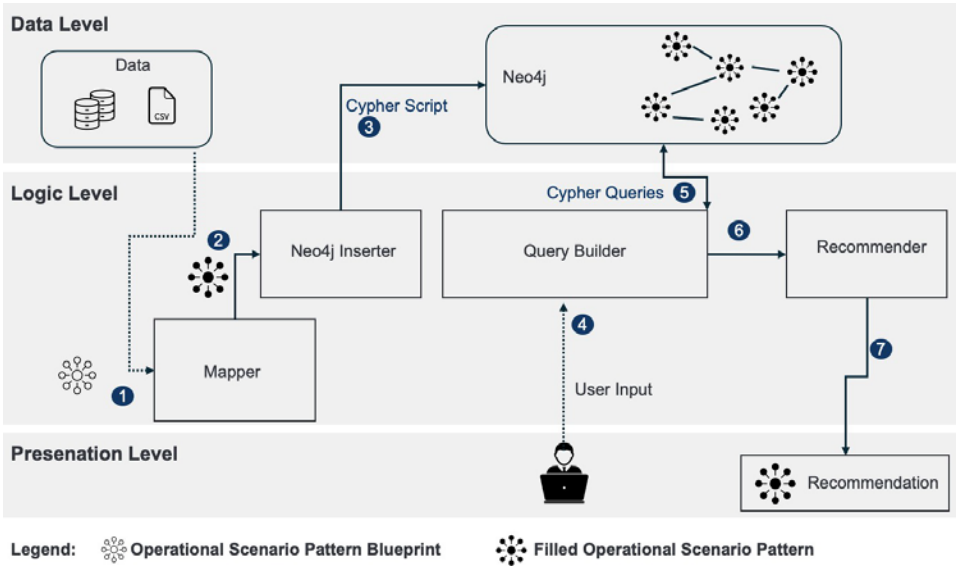


Figure 1. Graph-based Approach for Operations Planning in Civil Protection (GRETA)

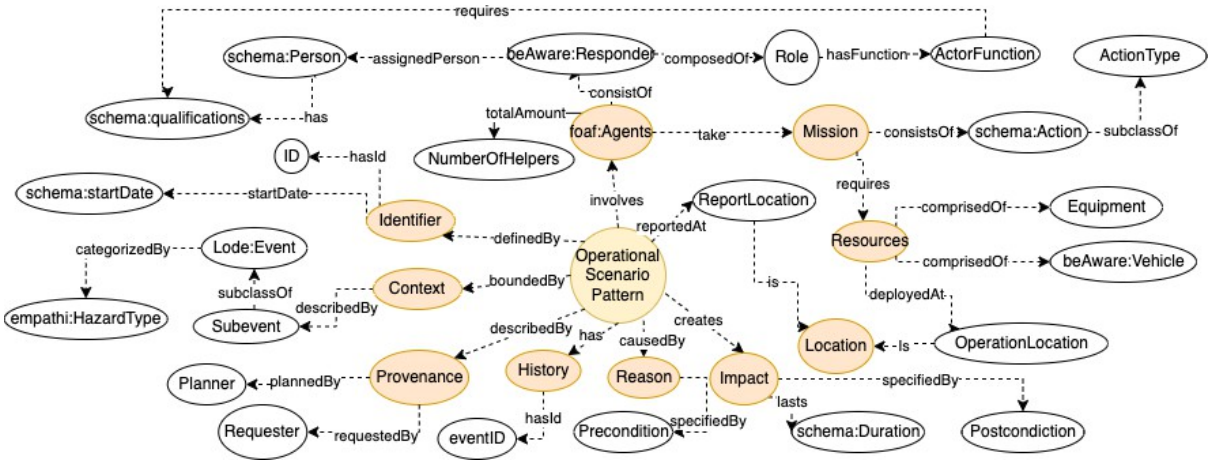


Figure 2. Graph-based view of Operational Scenario Pattern (Extract)

and assigned helpers per role (cf. Fig. 2). Furthermore, actions to be fulfilled in the operation are captured by the **missions** entity in form of action types as well as action sets per type. Vehicles required in operations are covered by the entity **resources**. Last, based on event type and location of the crisis (context), operational scenario patterns are interlinked by their identifiers within the **history** entity. All entities are anchored in and mapped onto existing vocabularies and ontologies (e.g. foaf[3], beAware[4], LODE[5], Empathi[6], DCAT[7], schema.org[8]) (cf. Fig. 3). GRETA applies the operational scenario pattern blueprint for mapping historic data on past operations on a knowledge graph with episodic operational knowledge (cf. Fig. 1). First, pre-processed data of historic operations in csv-format are mapped onto an operational scenario pattern blueprint in JSON-LD format by the Mapper component (cf. Fig. 1, step 1). Thereby, filled operational scenario pattern, means interlinked historic data are generated that are pushed to the Neo4j Inserter (cf. Fig. 1, step 2). By means of a cypher script the filled operational scenario patterns are inserted into Neo4j as knowledge base repository (cf. Fig. 1, step 3). Using JSON-LD as format has several advantages. First, the data entered into the knowledge graph can be easily expanded using other available data sets described by existing ontological standards. In addition, data of civil protection organizations such as the German Federal Agency for Technical Relief (THW) is currently only available in CSV format. However, CSV is not suitable

3 http://xmlns.com/foaf/0.1/
4 https://github.com/beAWARE-project/ontology
5 https://linkedevents.org/ontology
6 https://w3id.org/empathi/1.0
7 http://www.w3.org/ns/dcat
8 http://schema.org

Entities	Attributes	THW Data	Ontology Mapping
Identifier	ID, start date, end date	ID, start, end	-, schema:startDate, schema:endDate
Context	Event type, event, description	Event type, event, description of event	empathi:HazardType, Lode:Event, schema:description
Provenance	Planning Organization, Requesting Organization	Planner & requester	schema:Organization
Location	Regional area, federal state, city, country	Regional area, federal state, city, country	schema:addressRegion, schema:State, schema:addressLocality, schema:addressCountry
Reason	Preconditions	Service project	-
Impact	Duration, Postconditions, severity	Planned duration, report of operation	schema:duration, empathi:Severity
Actors	Total number of actors, workhours, unit, strength, role, helpers per role	Number of helpers, workhours, units, strength, function, Number of helpers needed per functior	foaf:Agent,-, -,beAWARE:Responder, -, -
Mission	Action type, action set	Task, task description	beAWARE: Mission,-, schema:action
Resources	Vehicles	Vehicles	beAWARE: Vehicle
History	Scenario ID	Historical IDs	-

Figure 3. Operational Scenario Pattern

for mapping nested tree-like structures as presented by civil protection units. JSON-LD enables to keep important structural and semantic information between units, functions, their tasks to be performed and required qualifications that are transferred to the knowledge graph. This in turn enables further completion of meanwhile unconnected individual datasets through inferencing and a deeper analysis of the data records with regard to graph metrics, e.g. to recognize key personnel via the number of connected nodes and edges for certain events. The episodic representation of operational knowledge can then be requested using cypher queries by means of the Query Builder (cf. Fig. 1, step 4 and 5). To give an example, the following query extracts information on actors connected to a specified event: `MATCH (osp:OperationalScenarioPattern)-[:t_bounded_by]-(c:Context)-[:t_described_by]-(s:Subevent) WHERE s.name =~ "(?i).*$SUBEVENT_NAME.*") WITH sp MATCH (sp)-[:t_involves]-(i:Actors) RETURN distinct(i)`.The resulting output is processed by the Query Builder and aggregated to planning recommendations by the Recommender (cf. Fig. 1, step 6 and 7). Recommendations are structured according to a subset of operational scenario patterns; here the entities Actors and Impact are used to present results (cf. Fig. 4). GRETA was implemented as a web-based service using HTML, JavaScript, JSON-LD, Neo4j[9] [10].

USE CASE

We instantiated and evaluated GRETA in a concrete use case with the German Federal Agency for Technical Relief (THW) - an agency of the German Federal Ministry of Interior and Community. THW operations include for instance debris clearance, electricity supply, bridge building or logistics, serving the purpose of maintaining public structures and critical infrastructure. Currently, the operations planning process at THW is deployed manually by experts. Although, THW collects data on larger missions, data about included operations differ with respect to formats and details of fulfilled operation tasks etc. Up til now, available data are therefore of limited usefulness for planning experts when allocating tasks and personnel to upcoming operations. Furthermore, structured data on historic operations including allocated units and tasks are not accessible at all via the internal IT infrastructure. This information is embedded in heterogeneous, unstructured textual documentations that do not offer direct added value in the effective and efficient planning of operations under time pressure of an expected crisis event. Within the use case, a data set covering #157,450 historic THW operations within Germany (2012-2022) was used as input for GRETA, including details on crisis events with respect to location and type, requester and planner of operation, date, planned duration, report on tasks, number of helpers etc.

[9] https://github.com/InformationServiceSystems/pairs-project/blob/main/Modules/THW
[10] https://www.youtube.com/watch?v=ohWVGATnjZM&t=1s

Entities	Attributes	Recommendation
Actor	TotalNumberOfHelpers	12
		135
	Responders	'Debris Clearance', 'Technical Unit Waterdamage/Pumping', 'Diving Group', 'Bridge Construction Group', 'Emergency supply and repair', 'Locating group', 'Technical Platoon'
Impact	Duration	0:10:32:08

Figure 4. Recommendation for operation plan for upcoming crisis event based on user input "flood" as *event type* and "secure operation site" as *action type*

Data Preparation

In order to enable the mapping of the historic operations data to operational scenarios patterns (cf. Fig. 2), the data set was prepared, i.e., we (1) collected, (2) cleaned, and (3) merged the data entries to generate a comprised data set.

(1) Collecting data: Initially, the data set included details on the crisis event such as location and type, the fulfilled operation, requester and planner, date, planned duration, a short report (description), number of deployed helpers, internal operation ID, provided service and assigned service project. For adding missing information on involved units and vehicles of the historic operations, an available tool developed by THW was used for extraction[11]. The extracted data covered details about tasks associated with over 40 units and details on 57 vehicle types. Additionally, 47 strength and equipment certification documents (StAN) were transformed into a structured data set for extending the data set with information on assigned roles and vehicles per unit.
(2) Cleaning data: The resulting data set was cleaned by deleting rows with missing features, merging operations with the same internal ID and correcting features with deviant formats and broken entries (e.g., dates, planned duration). As a result, the cleaned data set covered #31,118 operations.
(3) Merging data: In order to be able to inference on deployed units, tasks and vehicles per historic operation, collected data sets were merged using the BERT model (Reimers and Gurevych, 2019). Embeddings of operation descriptions and tasks associated with units as well as vehicles were generated. Furthermore, cosine similarity was used to establish a relationship between operations units based on unit's potential tasks and the tasks conducted within the operations. By testing diverse thresholds, we found an acceptable trade-off between the number of data entries and an accurate allocation of units per operation (confidence score: 60%). As a result, the integrated data set consisted of #15,274 historic operations.

THW Operations Planning Supported by GRETA

The pre-processed data set of #15,274 THW operations was mapped onto the operational scenario pattern blueprint in JSON-LD by the Mapper component of GRETA. Fig. 3 shows the mapping relations between entities of the pattern to THW data features. Resulting filled operational scenario patterns, i.e., historic THW operations are generated, pushed to the Neo4j Inserter and inserted into a knowledge graph (cf. Fig. 1). Giving an exemplary user request for planning an operation of the *event type:"flood"* and the *action type:"secure operation site"*, this input is translated into a cypher query by means of the Query Builder. The resulting output covers #306 relevant historic operations that are processed by the Query Builder, e.g., for inferencing the appropriate number of helpers, planned duration for the operation to be planned based on the historic data. The extracted data are aggregated to an operations plan recommendation based on the structure of the operations scenario pattern by the Recommender component, i.e., the entities actors and impact are used to present recommendation results (cf. Fig. 4). Given the exemplary user input, GRETA recommended to staff the planned operation with 12 helpers with 135 hours potential working time based on an expected operation duration of 10 hours and 32 minutes. Furthermore, units are listed that are required for conducting the operation, e.g., debris clearance, and bridge construction group (cf. Fig. 4). Generated results are represented within a first prototypical service, that is available as open source code on Github[12]. The user interface comprises four tabs (cf. Fig. 5). Two tabs are used to describe the service and applied Operational Scenario Patterns (1) Introduction & Data, (2) Background Information, while the remaining two tabs serve as the main functionalities of the service (3) Event Prediction, (4) Recommendations. Based on the amount of data per individual crisis event, that was available to train our predictive machine learning forecasting model, the event prediction focuses on weather-based events, in particular floods, heavy rainfall and snowfall within Germany.

[11] https://gitlab.com/Manuel_Raven/dienststellen-scraper-extended
[12] https://github.com/InformationServiceSystems/pairs-project/blob/main/Modules/THW



Figure 5. Overview of the user interface for the implemented web service of GRETA

After selecting the desired location, the user receives information on the occurrence (red dots) of predicted events for the next 7 days, as well as a risk assessment for the forecast via the confidence of the forecasting model and a threshold value (orange line). In addition, a map-based area overview provides topographical information and current weather information on selected locations. Planning recommendations are generated for the selected location and predicted crisis events and an additionally selected task category. The recommendations include a classification of the event, contextual information on similar past incidents, recommendations for action, personnel recommendations and resource recommendations per unit. A demo of the service is also available on Youtube [13]. Within the case study, we showed the potential of GRETA for supporting operations planning in civil protection in terms of required time (efficiency), accuracy and completeness of resulting plans (effectiveness) as well as the empowerment of less-experienced planning officers. In contrast to actual dispersed data and isolated documents at THW, operational scenario patterns supports structuring available data of past operations by means of a linked data approach (JSON-LD) and improves the accessibility of existing operational knowledge. Furthermore, GRETA offers the opportunity to interlink operational knowledge by diverse civil protection organization by applying operational scenario patterns on their data enabling a cross-organizational operational knowledge space for more efficient collaboration in operations planning and conduction.

LIMITATIONS AND FUTURE WORK

While GRETA demonstrates promising results in enhancing operations planning for civil protection organizations, there are also several limitations. The effectiveness of GRETA heavily relies on the quality and completeness of historical data. Incomplete or inaccurate data may lead to suboptimal recommendations. Future work should focus on improving the data collection methods and ensuring data accuracy to enhance the reliability of GRETA (cf. section on data preparation for the use case). The scalability of GRETA in handling larger datasets and diverse crisis scenarios needs further investigation. As the volume and complexity of crisis events increase, the scalability of GRETA's knowledge graph construction and recommendation generation processes may become a bottleneck. Future research should explore strategies for optimizing GRETA's performance on larger datasets also with respect to cross-organizational operation planning based on shared data on historic operations.

While GRETA utilizes operational scenario patterns to represent domain knowledge, there may still be nuances specific to civil protection operations that are currently not adequately captured. Enhancing the domain-specific knowledge representation within GRETA could lead to more accurate and tailored recommendations for operations planning. Those aspects will be investigated in an empirical user study with the resulting prototype with experienced planning officers by THW. In this study also the usability and adoption of GRETA by planning officers will be evaluated in real-world settings as those are crucial factors for its practical utility.

Furthermore, GRETA primarily relies on historic data for generating operation plans. Integrating real-time data streams, such as weather data, social media feeds or sensor data (e.g., water level of rivers), could further enhance GRETA's capabilities by providing up-to-date information on unfolding crisis events and potentially required operations. Future research will explore methods for integrating real-time data and generate predictions for potential crisis events by applying machine learning methods (e.g., XGBoost) into the decision support system based on GRETA.

[13] https://www.youtube.com/watch?v=ohWVGATnjZM&t=1s

CONCLUSION

In this paper, we introduced GRETA, a graph-based approach for enhancing operations planning in civil protection organizations. By leveraging semantically enhanced historic data and operational scenario patterns with linked data, GRETA provides decision support for generating operation plans for upcoming crisis events. Through a case study with the German Federal Agency for Technical Relief (THW), we demonstrated the potential of GRETA in improving the efficiency, accuracy, and completeness of operation plans, while empowering especially less-experienced planning officers.

REFERENCES

Ali, S. F. (2015). Towards peer pressure in post-disaster governance: An empirical study. *Hastings International & Comparative Review*, *38*, 243.

Bompotas, A., Anagnostopoulos, C., Kalogeras, A., Kalogeras, G., & et al. (2022). A civil protection early warning system to improve the resilience of adriatic-ionian territories to natural and man-made risk. *In 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*, 1–8. https://doi.org/10.1109/ETFA52439.2022.9921697

Chehade, S., Matta, N., Pothin, J.-B., & Cogranne, R. (2018). Data interpretation support in rescue operations: Application for french firefighters. *In 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, 1–6. https://doi.org/10.1109/AICCSA.2018.8612779

Correia, A., Severino, I., Nunes, I. L., & Simo˜es-Marques, M. (2018). Knowledge management in the development of an intelligent system to support emergency response. *Proceedings of the AHFE 2017 International Conference on Human Factors and Systems Interaction (AHFE)*. https://doi.org/10.1007/978-3-319-60366-7 11

Gaur, M., Shekarpour, S., Gyrard, A., & Sheth, A. (2019). Empathi: An ontology for emergency managing and planning about hazard crisis. *In 2019 IEEE 13th International Conference on Semantic Computing (ICSC)*, 396–403. https://doi.org/10.1109/ICOSC.2019.8665539

Habib, L., Pacaux-Lemoine, M.-P., & Millot, P. (2018). Human-robots team cooperation in crisis management mission. *In 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 3219–3224. https://doi.org/10.1109/SMC.2018.00545

Hassan, M. K. A., & Chen-Burger, Y.-H. (2016). Communication and tracking ontology development for civilians earthquake disaster assistance. *In 13th International Conference on Information Systems for Crisis Response and Management 2016 (ISCRAM)*.

Idris, A., & Soh, S. N. C. (2014). The relative effects of logistics, coordination and human resource on humanitarian aid and disaster relief mission performance. *South East Asian Journal of Management (SEAM)*, *8*(2), 1.

Janzen, S., Gdanitz, N., Abdel Khaliq, L., Munir, T., Franzius, C., & Maass, W. (2023). Anticipating energy-driven crises in process industry by ai-based scenario planning. *Proceedings of the 56th Hawaii International Conference on System Sciences (HICSS)*.

Kontopoulos, E., Mitzias, P., Moßgraber, J., Hertweck, P., & et al. (2018). Ontology-based representation of crisis management procedures for climate events. *In 15th International Conference on Information Systems for Crisis Response and Management 2018 (ISCRAM)*.

Krstikj, A., Esparza, M. G. C. R., Vargas, J. M., Escobar, L. H., & et al. (2021). Decision-support tool for coordination of volunteers in large-scale lockdowns. *International Journal of Disaster Risk Reduction*, *62*. https://doi.org/10.1016/j.ijdrr.2021.102420

Lee, E. K., Pietz, F. H., Chen, C.-H., & Liu, Y. (2017). An interactive web-based decision support system for mass dispensing, emergency preparedness, and biosurveillance. *Proceedings of the 7th International Conference on Digital Health*, 137–146. https://doi.org/10.1145/3079452.3079473

Lim, M. H., Leong, Y.-K., Tan, N. J. Y., & Ho, L.-Y. (2017). Complex scenario planning for disaster relief. *Proceedings of the 2017 International Conference on Intelligent Systems, Metaheuristics Swarm Intelligence (ISMI)*, 147–151. https://doi.org/10.1145/3059336.3059355

Lindner, S., Kuehnel, S., Betke, H., & Sackmann, S. (2018). Simulating spontaneous volunteers-a conceptual model. *In 15th International Conference on Information Systems for Crisis Response and Management 2018 (ISCRAM)*.

Mescherin, S. A., Kirillov, I., & Klimenko, S. (2013). Ontology of emergency shared situation awareness and crisis interoperability. *In 11th International Conference on Cyberworlds 2013 (CW)*, 159–162. https://doi.org/10.1109/CW.2013.61

Rauchecker, G., & Schryen, G. (2018). Decision support for the optimal coordination of spontaneous volunteers in disaster relief. *In 15th International Conference on Information Systems for Crisis Response and Management 2018 (ISCRAM)*. https://doi.org/10.5283/epub.37001

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Shittu, E., Parker, G., & Mock, N. (2018). Improving communication resilience for effective disaster relief operations. *Environment Systems and Decisions*, *38*, 379–397. https://doi.org/10.1007/s10669-018-9694-5

Wrona, K., Tortonesi, M., Marks, M., & Suri, N. (2019). Leveraging and fusing civil and military sensors to support disaster relief operations in smart environments. *In 2019 IEEE Military Communications Conference (MILCOM)*. https://doi.org/10.1109/MILCOM47813.2019.9021004

Zhao, X., Yuan, B., Sun, G., & Liu, P. (2022). Chinese semantic description framework of operational planning statements. *In 2022 8th International Conference on Systems and Informatics (ICSAI)*, 1–6. https://doi.org/10.1109/ICSAI57119.2022.10005359

Adapting PLUM: Earthquake Early Warning with Node-Level Processing in New Zealand

Chanthujan Chandrakumar
Joint Centre for Disaster Research
Massey University
cchandra2@massey.ac.nz

Marion Lara Tan
Joint Centre for Disaster Research
Massey University
m.l.tan@massey.ac.nz

Caroline Holden
SeismoCity Ltd., Wellington
caroline.francoisholden@gmail.com

Max T. Stephens
Civil and Environmental Engineering
University of Auckland
max.stephens@auckland.ac.nz

Raj Prasanna
Joint Centre for Disaster Research
Massey University
r.prasanna@massey.ac.nz

ABSTRACT:

Is running the Propagation of Local Undamped Motion (PLUM) algorithm in a community-engaged earthquake early warning (EEW) network feasible, and can it function effectively at the node level without centralised processing units? This study investigates the practicality of deploying the PLUM algorithm within a node-level architecture, shifting away from traditional centralised seismic data processing methods. The study uses cost-effective MEMS-based seismographs to decentralise EEW. The preliminary phase of the research included the deployment of sensors and the establishment of a two-tiered Primary-Secondary node structure for node-level intensity prediction and alert generation, with the sensors functioning as independent prediction points. Future work includes threshold calibration for optimal alert issuance, and network expansion to reduce blind spots. This work-in-progress paper discusses progress towards a scalable, efficient EEW system that could serve as a replicable model for earthquake-prone regions globally, aiming for operational readiness that empowers communities against the threat of earthquakes.

Keywords

Earthquake Early Warning (EEW), Node-Level Processing, PLUM Algorithm, Citizen Seismology, Low-cost sensors

INTRODUCTION

Earthquake Early Warning System (EEWS) has been recognised as a vital advancement towards mitigating the impacts of seismic events (Allen & Melgar, 2019; Chandrakumar et al., 2022). The evolution of EEWS has been propelled by technological advancements and research breakthroughs, resulting in the implementation of systems across various regions and countries (Chandrakumar et al., 2022; Cremen & Galasso, 2020; McBride et al., 2022). These systems are designed to provide advanced alerts, enabling individuals and authorities to take preventive measures to safeguard lives and infrastructure. The effectiveness of EEWS in reducing damages and injuries associated with seismic activities has been well documented globally (e.g., Fujinawa & Noda, 2013; Nakayachi et al., 2019; Suárez et al., 2009).

Despite the pronounced seismic hazard in Aotearoa New Zealand (NZ), the country lacks an official national EEWS, a critical gap in preparedness underscored by Becker et al. in their 2020 study. The country relies predominantly on the GeoNet programme, managed by GNS Science, for rapid earthquake information. GeoNet provides earthquake source parameters and ground-shaking data through a website and mobile app but lacks dedicated Earthquake Early Warning (EEW) capabilities (GeoNet, 2017).

The absence of an official EEWS in NZ is primarily due to various technical and non-technical challenges, with the high cost of advanced systems representing a considerable obstacle (Brooks et al., 2021; Prasanna et al., 2022). For cost-effective earthquake preparedness solutions, there has been a marked global trend towards adopting low-cost alternatives. Micro-electromechanical systems (MEMS) based ground motion sensors have gained prominence among these. Demonstrating a successful track record in seismic applications since the early 1990s, MEMS-based ground motion sensors have been effectively integrated into EEWS in various regions, facilitating real-time public alerting systems (Anthony et al., 2019; Holland, 2003).

In pursuing a cost-effective EEWS in NZ, Prasanna et al. (2022) have initiated and deployed an experimental, community-engaged network in the Greater Wellington region. As depicted in Figure 1, this innovative system leverages Raspberry Shake 4D (RS4D) seismographs equipped with MEMS-based accelerometers, signifying a pivotal development in earthquake detection and alert technology. This network adopts a node-level processing approach to analyse and process ground motion data, diverging from the centralised processing approaches predominant in established EEWS globally. The experimental network uses the propagation of local undamped motion (PLUM) algorithm (Kodera et al., 2018) for its EEW capabilities, optimised for low-cost sensors with limited computational resources, to facilitate this node-level processing.

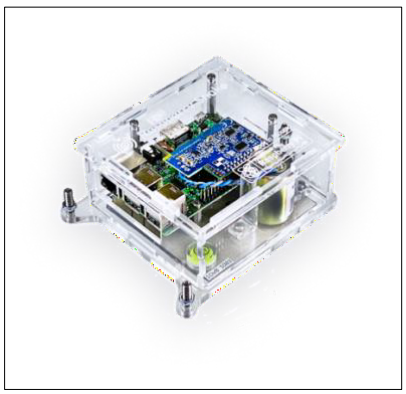


Figure 1: The RS4D Seismograph.

The effectiveness of node-level processing has been compared to conventional centralised approaches. This assessment was conducted by applying the PLUM algorithm to a selected array of sensors within a simulated environment (Prasanna et al., 2022). Chandrakumar et al. (2023) have also validated this approach's earthquake detection capabilities within a community-engaged network, focusing on a particular earthquake event. This work-in-progress paper builds on past research and addresses the ongoing progression and practical application of the PLUM algorithm within a real-life operational context, including adaptation processes of the original PLUM algorithm for node-level processing.

The article is structured as follows: The Background section outlines the node-level processing architecture and PLUM algorithm. The Work-in-Progress section details the ongoing adaptation of the PLUM algorithm to a node-level processing architecture within the community-engaged EEWS. Finally, the Discussion section presents the insights from the preliminary findings, outlines future research work, and concludes by highlighting the importance and potential impact of this research.

BACKGROUND

Node-Level Processing Architecture

Two principal methodologies dominate the landscape in data processing: centralised and node-level processing (Xi, 2020). Centralised processing gathers data from diverse sources into a unified, centralised facility undergoing analysis and processing. Conversely, node-level processing adopts a decentralised approach, wherein individual nodes or sensors independently execute data analysis and processing tasks (Santamaria et al., 2019). This distinction primarily revolves around allocating computational responsibilities; centralised systems centralise data processing in a singular hub, whereas node-level systems distribute these tasks across each node within the

network, potentially augmenting the system's responsiveness and robustness against widespread failures.

Most EEWS operates on a centralised model, where a central facility receives and processes seismic data. While this centralised approach offers superior control and consistency in data handling, it is susceptible to critical delays and communication failures, particularly when major seismic events compromise infrastructure (Prasanna et al., 2022).

Recent technological advances have led to the enhancement of low-cost, MEMS-based sensors with in-built processing capabilities. These advancements enable data processing directly at the sensor nodes, allowing a transition from centralised to node-level processing (Chandrakumar et al., 2022). Node-level processing empowers the system to analyse seismic information and issue alerts where the data is collected, utilising the built-in computational power of the sensors. Aside from Prasanna et al.' (2022) publication on community-engaged EEW network, there is minimal documentation of node-level processing for EEW.

The shift toward node-level processing is motivated by multiple benefits. It reduces the time to issue alerts by removing the necessity for data to travel to a central server, allowing the system to provide a longer warning window for the end users. It also strengthens the resilience of the EEWS against central system failures caused by earthquakes. Furthermore, processing at the node level can decrease the operational costs associated with the upkeep of sophisticated central processing systems (Fischer et al., 2012; Prasanna et al., 2022).

The PLUM Algorithm

Over the past three decades, significant advancements have been made in EEWSs (Chen et al., 2015; Clinton et al., 2016). These advancements have primarily focused on network-based[1] approaches, utilising source-based[2] models to estimate earthquake parameters like hypocentre locations and magnitudes (Kuyuk et al., 2014; Satriano et al., 2011). Traditional EEW systems analyse seismic data, including P-wave onsets and maximum displacement amplitudes (Meier et al., 2015; Noda et al., 2016). However, the accuracy of these predictions heavily depends on the precision of the estimated parameters, which poses challenges, particularly during complex seismic events like large magnitude earthquakes and simultaneous aftershocks (Hoshiba, 2013; Hoshiba et al., 2011).

Recognising these limitations, Kodera et al. (2018) introduced an alternative approach known as the PLUM method. Based on Hoshiba (2013) work, this method offers a more direct ground motion estimation approach, emphasising cost-effectiveness and ease of implementation.

The PLUM algorithm operates by forecasting seismic intensity using real-time ground motion data. Its operation involves collecting pseudo-seismic intensities from observation stations within a predefined radius around a target station. The forecasted seismic intensity is then determined based on the highest value among these real-time observations. This method assumes that ground motion, responsible for significant seismic intensity, travels within the specified radius without notable attenuation. Figure 2 illustrates the PLUM algorithm's operation, demonstrating the intensity prediction at a prediction point using intensity data from surrounding observation stations.

[1] A network-based approach in EEWS involves deploying a network of sensors across various locations within a geographical area. Earthquake detection is achieved by collectively processing the data gathered from this interconnected sensor network.

[2] Source-based EEWS models are designed to detect earthquakes and alert stakeholders, providing detailed information about the seismic event.

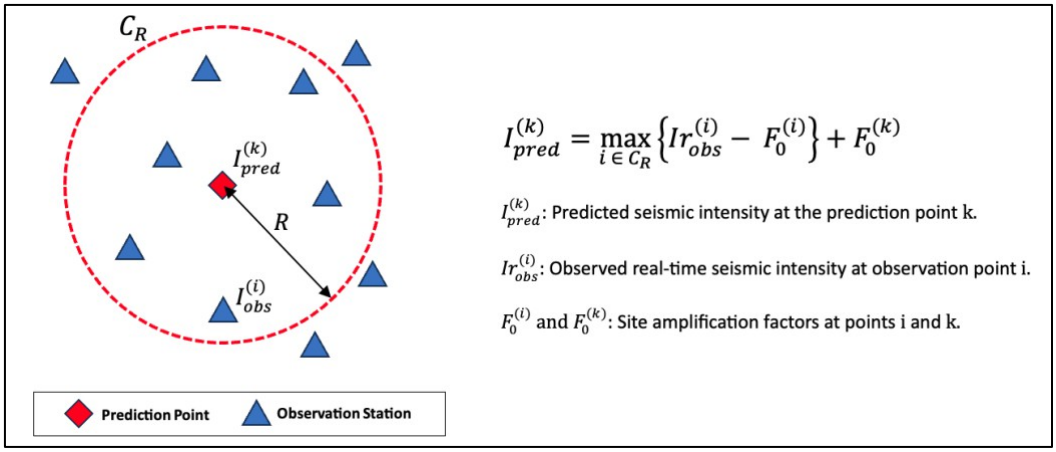


Figure 2: Illustration of the PLUM Algorithm's Operation, Demonstrating Intensity Prediction at a Target Station Based on Observations from Peripheral Stations (Adapted from Kodera, 2018).

A key component in the operation of the PLUM method is monitoring the observed intensity at each observation station. The PLUM method triggers an EEW alert when the real-time ground motion predictions from two or more monitoring stations exceed a specified threshold. This alert is generated explicitly for the region surrounding the target station, ensuring timely warning dissemination.

The PLUM method's distinct advantage lies in its ability to continuously monitor ground motion distribution in real-time, addressing the overprediction issues and false alarms prevalent in traditional ground-motion-based systems. This was particularly evident following the Mw 9.0 Tohoku earthquake, where PLUM issued more reliable early warnings (Hoshiba, 2021). Further validation with the 2016 Kumamoto earthquake data reinforced PLUM's superiority, consistently providing earlier warnings and effectively reducing blind zones (Kodera et al., 2016).

Moreover, the PLUM method has shown resilience in handling multiple simultaneous earthquakes and predicting strong motion by monitoring real-time seismic intensities. This capability sidesteps technical issues associated with source-based models and suggests a potential for further improvements with denser seismic networks, including low-cost sensors (Kodera et al., 2016).

Despite its advantages, the PLUM method has limitations. Its effectiveness is contingent upon a dense observational network, as it is limited by the short prediction radius and site amplification factors that significantly influence seismic ground motion (Kodera et al., 2016). Additionally, the PLUM algorithm tends to provide shorter warning durations than source-based methods. While it excels in predicting ground motion in complex seismic scenarios, the maximum warning time it offers is limited by the radius of observation.

WORK IN PROGRESS

This section highlights the work-in-progress of this study's implementation of an EEWS with node-level processing for the PLUM algorithm.

Installation of Seismic Sensors and Validation of Ground Motion Data

The initial phase of seismic sensor installation was carried out in 2020 and 2021, utilising MEMS-based RS4D sensors for their affordability, accuracy, and processing power, aligning with the objectives of a community-engaged EEWS (Prasanna et al., 2022). Community participants were recruited to host these ground motion sensors within their households. An workshop was conducted to gain insights into user perspectives and to develop strategies for distributing sensors to the general public, promoting a citizen seismology approach (Tan et al., 2021).

Following that, the network expansion in 2023 aimed to increase sensor density, leveraging social media campaigns to gather community interest in hosting additional RS4D sensors. The campaign resulted in 80 expressions of interest, informing the strategic selection of sensor locations through geographical data analysis. By mapping the geographical spread of these responses, 25 locations were identified from the pool to install additional sensors. This expansion strategy aligns with insights from a recent study, which explores the necessity of a comprehensive approach to developing a community-engaged EEWS for NZ, addressing the technological, societal, and cultural challenges (Tan et al., 2023). The criteria for site selection encompassed algorithm coverage, seismic activity, proximity to fault lines, population density, and other technical and environmental considerations. This careful selection has contributed to a network comprising approximately 60 ground motion

sensors across NZ, with a significant concentration of around 45 sensors within the Greater Wellington area.

The performance of the community-engaged EEW network was evaluated through a dedicated study focusing on P-wave detection capabilities utilising the collected ground motion data (Chandrakumar et al., 2023). This evaluation centred on data obtained within 48 hours surrounding a magnitude 5.8 earthquake on 22 September 2022. The analysis involved a detailed assessment of recorded seismic data by the stations, specifically focusing on identifying any false detections or missed detections of seismic events within the dataset. The findings from this investigation underscore the effectiveness of a citizen seismology-based EEWS in identifying seismic occurrences relevant to early warning objectives.

These results affirm the operational capability of the newly implemented network, demonstrating its potential to contribute meaningfully to earthquake early warning efforts.

Adaptation of PLUM for Node-Level Architecture

Reason for Choosing PLUM for Node-level Processing.

In developing the node-level architecture for EEWS, the crucial task was to select an algorithm that conforms to the computational limits of the RS4D sensor nodes. These nodes are limited by their processing capabilities, necessitating an EEW algorithm that is both lightweight and effective. Through benchmarking existing EEW algorithms, the PLUM algorithm stood out as the optimal solution with its robustness, lightweight design, and ease of implementation (Prasanna et al., 2022). Its successful deployment in Japan's EEWS and validation across various global contexts affirm its appropriateness (Kodera et al., 2018). With its proven effectiveness with EEWSs worldwide, the PLUM algorithm was chosen for node-level operation for this study.

Processing Method of the Original PLUM Algorithm

The PLUM method predicts seismic intensities at designated prediction points, distributed across various areas based on seismic risk. For each area, the predicted seismic intensity is computed as the highest forecasted value among all its prediction points. This computation involves selecting the maximum observed real-time seismic intensity within a radius R from each prediction point, incorporating site effect corrections with a continuous one-second update interval (Kodera et al., 2018).

Real-time ground motion data processing in PLUM is centralised; observation stations transmit their data to a central unit, which issues intensity predictions to the target sites. This system has pre-programmed prediction points and corresponding observation stations for streamlined processing (Kodera et al., 2018). However, the adaptation to node-level processing necessitates a fundamental reconfiguration. In this architecture, each sensor node assumes the role of processing, effectively rejecting the need for a centralised processing unit. This shift ensures that the PLUM algorithm's calculations are conducted within the individual sensor nodes, thus fully embracing the decentralised nature of node-level architecture.

Defining Prediction Points for Node-level Architecture

The next objective focused on enabling the functionality of the PLUM algorithm at the node level. Defining prediction points is a foundational step in the methodology for adapting the PLUM algorithm to node-level architecture. The current implemented EEW network is depicted in Figure 3. In the original implementation of the PLUM method, authorities selected prediction points based on regional seismic activity to disseminate alerts to those areas (Kodera et al., 2018). These regions had various prediction points, reflecting the seismic activities observed.

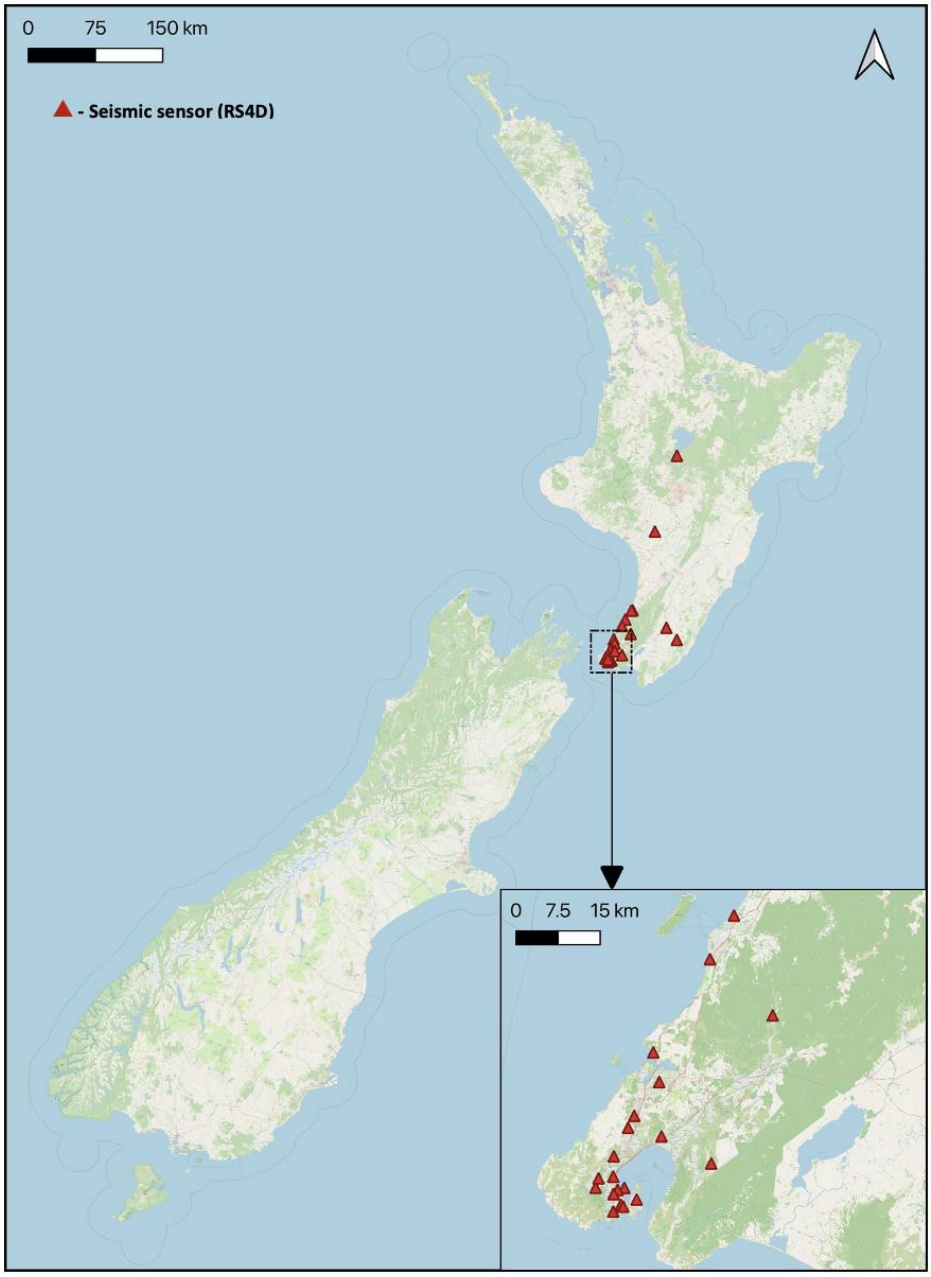


Figure 3: Overview of the EEW Network Deployed by the CRISiSLab Team.

Compared to the original PLUM implementation, our approach differs by designating the sensors as prediction points. Figure 4 presents four different prediction points as an example; each sensor marked in red represents a prediction point with a 30 km prediction radius consistent with the original PLUM specifications. Sensors within this radius serve as observation stations around the prediction point and are responsible for predicting the seismic intensity at the prediction point. Currently, our focus is not on issuing EEW alerts for broader regions but on generating alerts specifically for the prediction point - the RS4D sensor. This strategy is rooted in the concept of a community-engaged EEW network, ensuring that each sensor host receives EEW alerts directly through their installed RS4D sensor. However, it is essential to note that, in future, the network has the potential to be augmented with additional passive nodes. These nodes would also function as prediction points but receive alerts without estimating the intensity, enhancing the system's reach and effectiveness.

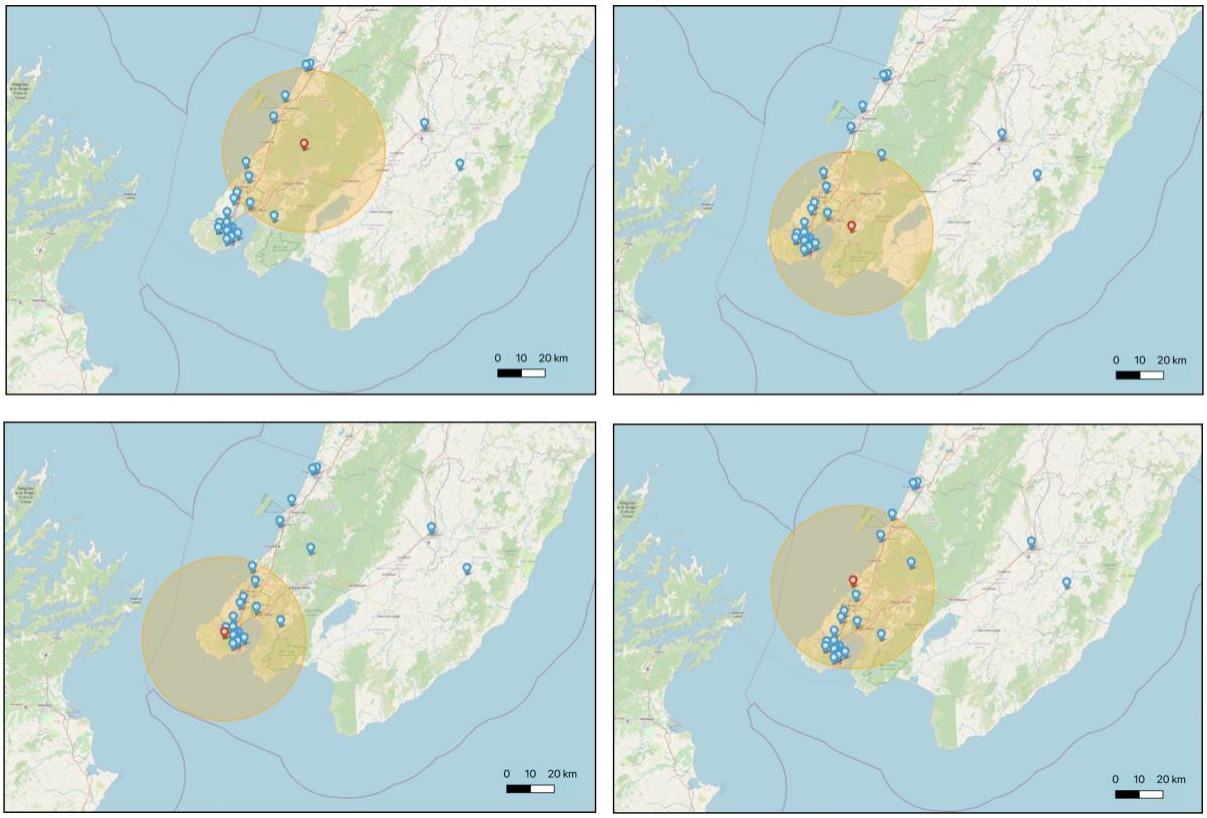


Figure 4: Illustration of Four Prediction Points with 30 km Prediction Radii, Highlighting Sensors in Red as Designated Prediction Points.

Implementing Node-Level Processing in the EEWS

Upon designating prediction points within the EEWS, the node-level implementation of the PLUM algorithm was initiated, establishing a dual-tiered structure of RS4D sensors: Primary and Secondary. Each prediction point, with a 30 km radius, is approached as a distinct scenario where one RS4D sensor is appointed as the Primary sensor, entrusted with two key functionalities: **intensity prediction** and **alert generation** for its specific prediction site. Concurrently, the surrounding sensors, classified as Secondary sensors within the same radius, are tasked with transmitting real-time intensity data to the Primary sensor for analysis. It is essential to highlight that a single sensor may serve multiple roles, acting as a Primary sensor for one prediction point while simultaneously functioning as a Secondary sensor for another, especially in areas where prediction radii overlap. The selection of a Primary sensor is strategically based on its current processing load, favouring sensors with minimal existing commitments to other prediction points. This strategic allocation and distribution of processing responsibilities ensure the efficiency and effectiveness of the node-level processing within our EEWS. Figure 5 illustrates how node-level processing utilises the PLUM algorithm, showcasing a prediction radius (Orange circle) scenario and the prediction point (Red icon), Primary (Purple icon) and Secondary stations (Blue icons).

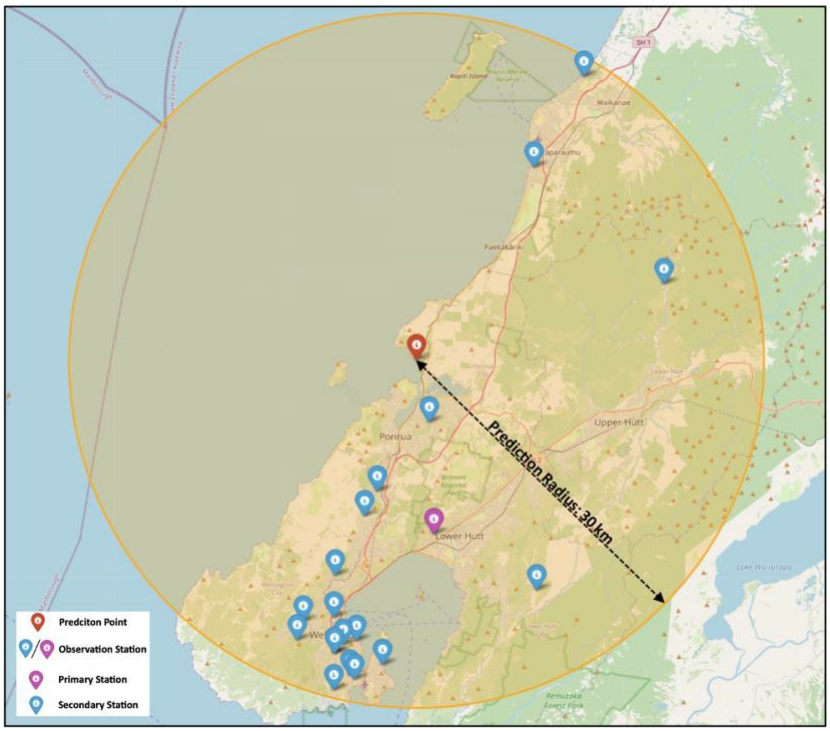


Figure 5: An Illustration of the PLUM Algorithm Adopted for Node-level processing, Featuring a Prediction Point, Primary Station, and Secondary Stations within a Specified Prediction Radius.

To address the potential failure of a Primary sensor, our system allocates at least three sensors to the Primary sensor role - one serves as the active Primary sensor, while the other two are on standby as backups. These backups also function as Secondary sensors, springing into action as the Primary sensor should the main one fail. This redundancy is crucial for maintaining continuous and uninterrupted seismic intensity predictions at each prediction point.

In contrast to the original PLUM methodology, which requires that all observation stations within the prediction radius update their real-time intensity to a centralised processing unit each second, our approach modifies the interaction between Secondary and Primary sensors. In our system, Secondary sensors independently process initial data, forwarding intensity updates to the Primary sensor only upon surpassing an internal threshold. Each sensor within our network is programmed to estimate real-time intensity, including Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV), and Peak Ground Displacement (PGD) on a second-by-second basis, assessing whether the PGA exceeds a specified internal threshold.

The Primary sensor undertakes the task of predicting the intensity at the prediction point solely under two conditions: if its internal threshold is exceeded or if it receives an intensity update from any Secondary sensor within the prediction radius that has surpassed the internal threshold. Secondary sensors are programmed to transmit real-time intensity data (PGA, PGV, and PGD) to the Primary sensor only after crossing this pre-determined internal threshold; the specifics of determining the internal threshold are detailed in a separate subsection. This strategic introduction of an internal threshold diverges from the original PLUM's approach by aiming to minimise the processing load on the Primary sensor. By reducing the influx of data packets during PLUM operation, this internal threshold mechanism ensures the Primary sensor maintains optimal performance, marking a significant adaptation in our node-level implementation to streamline processing and enhance efficiency. Figures 6 illustrates the overall working flow of Primary sensor in terms of predicting the intensity to a given prediction point according to our PLUM approach adopted for node-level processing. The comprehensive operation of the Primary sensor's alert generation is detailed in the next section.

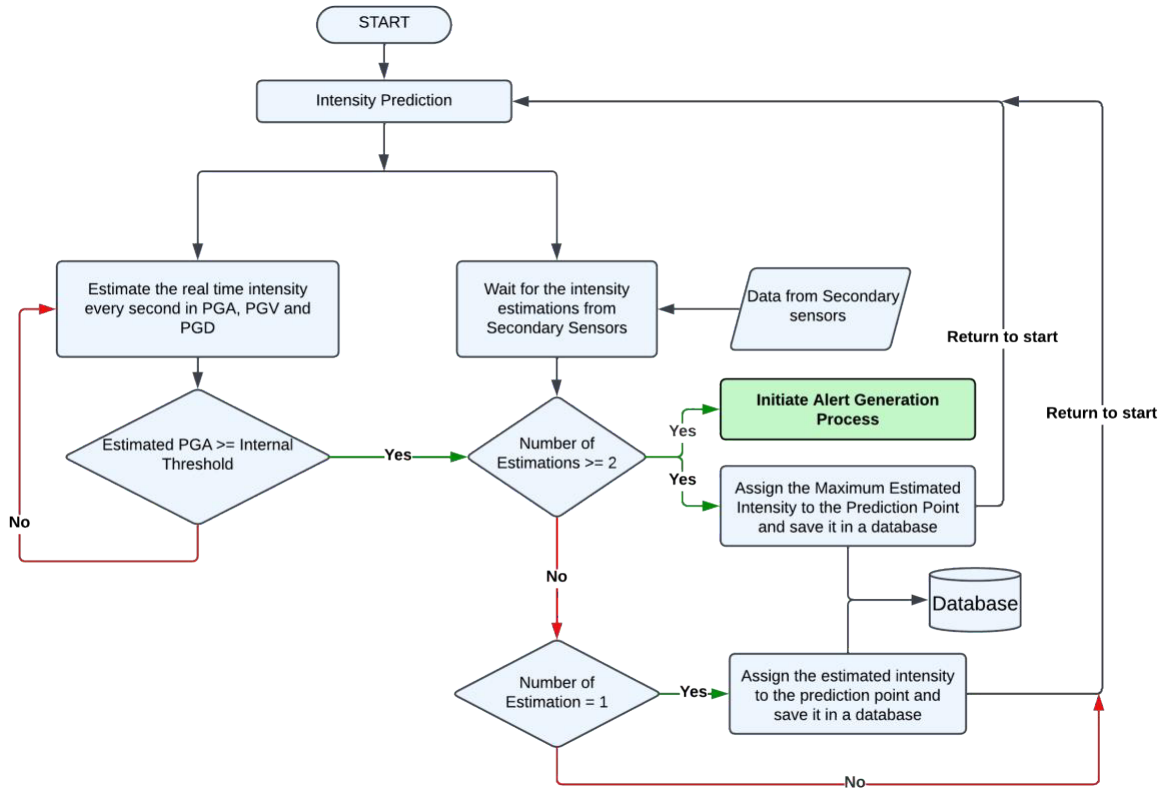


Figure 6: Flow Chart Explains the Intensity Prediction Flow of the Primary Sensor for the Assigned Prediction Point.

Defining the Internal Threshold for Secondary-Primary Communication

An internal threshold mechanism has been introduced to streamline communication between Secondary and Primary sensors, thereby reducing the processing load on the Primary sensor. This threshold determines when Secondary sensors should transmit real-time intensity updates. The formulation of this internal threshold began with establishing a specific Modified Mercalli Intensity[3] (MMI) range that effectively signals the necessity for an update without indicating considerable ground shaking.

The selection of an appropriate MMI range was based on the descriptions provided by GeoNet (Dowrick, 1996; GeoNet, n.d.). This choice ensures that our system's criteria for triggering intensity updates are well-tuned to the regional seismic activity profiles, thereby maintaining the system's accuracy and relevance. After careful consideration, MMI 3 was selected as the internal threshold. MMI 3 is characterised by indoor light vibrations, where hanging objects might sway slightly. This level of intensity serves as a pragmatic threshold, allowing Secondary sensors to initiate communication with the Primary sensor early in the event of significant seismic activity or at the onset of a large earthquake, thereby optimising the timing of alert dissemination.

Following the selection of MMI 3 as the internal threshold, the challenge was addressed by translating this MMI value into measurable parameters for the sensors, specifically PGA, PGV, and PGD. Since our sensors estimate intensity through these physical measures every second, determining an equivalent PGA value for MMI 3 was essential. To accomplish this, Ground Motion Intensity Conversion Equations (GMICEs), as outlined by Moratalla et al. (2020), were utilised. Figure 7 visually illustrates the correlation between MMI levels and PGA values, as defined by the GMICEs. In the figure, the MMI threshold of 3 or above is marked with a green vertical line, setting a clear benchmark for triggering communication from Secondary to Primary sensors based on real-time seismic data.

[3] The Modified Mercalli Intensity scale is a qualitative tool used to describe the intensity and effects of an earthquake on the Earth's surface, structures, and inhabitants. It ranges from I (imperceptible) to XII (destruction), describing the impact on structures, the environment, and human perception and response.

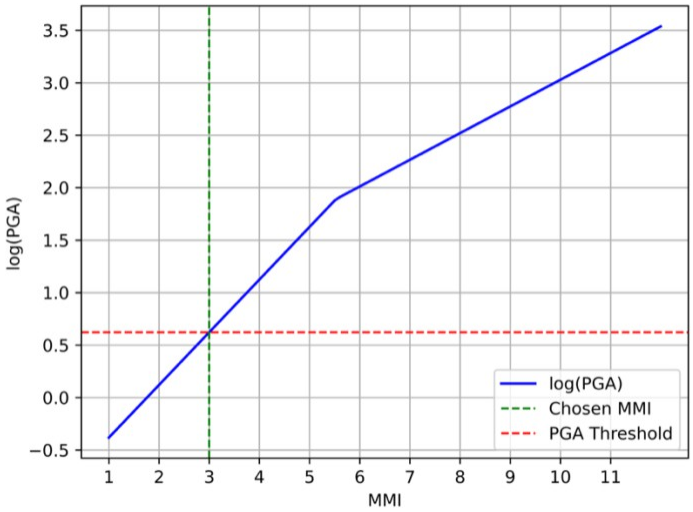


Figure 7: Correlation Between MMI Levels and PGA Values as Defined by GMICEs, Featuring a Green Vertical Line to Indicate the MMI Threshold of 3 and a Horizontal Red Dotted Line Representing the log(PGA) Threshold for Transferring Intensity Estimations from Secondary to Primary Sensor

Subsequently, the log(PGA) threshold was identified by analysing the intersection point depicted in Figure 7. The chosen log(PGA) value, serving as the internal threshold for Secondary sensors, was identified as 0.62, as represented by a red dashed horizontal line within the figure. After converting the logarithmic value to its original scale, the PGA internal threshold was determined to be 4.2 cm/s².

Alert Generation

In our EEWS, the generation of alerts for a specific prediction point, typically another RS4D sensor installed in a household, is managed by the Primary sensor designated for that point. The key requirement for a Primary sensor to trigger an alert is the reception of at least two intensity predictions that surpass a predefined threshold. These predictions can originate from the Primary sensor and one or more Secondary stations or exclusively from Secondary stations. The surpassing of this threshold by the predictions is critical, as it indicates a level of seismic intensity warranting generating an alert.

In the experimental network designed for testing, an intensity threshold of MMI level 5 has been provisionally selected. According to the GeoNet MMI descriptions, this intensity level corresponds to shaking typically felt outside, capable of awakening most sleepers and causing alarm among a few people indoors. However, establishing an ideal alert activation threshold requires carefully evaluating societal impacts and technical capabilities. Determining such a threshold is part of a broader, ongoing conversation to formulate a comprehensive set of guidelines specific to NZ's unique needs and circumstances.

The MMI 5 was translated into a PGA value of 42.3 cm/s² to quantify this threshold, utilising the same method for establishing the internal threshold for Primary-Secondary sensor communication. Therefore, if the predicted intensity from two or more observation stations within a prediction radius exceeds this PGA threshold of 42.3 cm/s^2, the Primary station will initiate an alert to the prediction point.

DISCUSSION

Launching a node-level EEWS using the PLUM algorithm in NZ seismically varied landscape represents a significant shift from conventional centralised models. Table 1 summarises the key differences between the traditional PLUM methodology and this research's approach. This proposed approach has optimised the processing capability of RS4D sensors for real-time seismic intensity prediction and alert dissemination. Transitioning to a decentralised architecture enhances the resilience of the EEWS by minimising the risks associated with a single point of failure, a potential issue in centralised systems.

Table 1: Comparative Overview of Traditional Plum Approach Versus Node-Level PLUM Adaptation Proposed in this Study.

Characteristics	Traditional PLUM implementation	Experimental Node-Level PLUM adaptation
Processing Approach	Centralised processing	Node-level processing with a two-tiered Primary-Secondary structure
Prediction Point Determination	Selected based on areas of interest reflecting observed seismic activities	Sensors themselves are used as prediction points
Data Handling Method	All observations within a prediction radius are sent to a centralised unit in real time.	Secondary sensors process data independently, sending updates to a Primary sensor only if internal threshold is surpassed.
Intensity Predicting Method	Observations are collected and processed at a central location to predict the intensity at prediction points.	The Primary sensor is responsible for predicting the intensity when its threshold is exceeded or upon receiving data from Secondary sensors.
Alert Generation	Centralised system issues alerts to the prediction points.	The Primary sensor issues an alert for its prediction point.

Adapting the PLUM algorithm for node-level application showcases the system's adaptability and the algorithm's suitability for environments with limited resources, particularly with the constraints of RS4D sensors. Identifying sensors as prediction points facilitates a community-centric warning system, ensuring swift alert delivery directly where needed. The classification of sensors into Primary and Secondary roles optimises task distribution, streamlining the network's operation.

A key focus for this proposed architecture was to minimise network traffic coming from the frequent intensity updates from Secondary sensors. Implementing an internal threshold curtails unnecessary data transfers, minimising the processing burden on Primary sensors. This threshold, determined by NZ's specific MMI scale and GMICEs, enhances network efficiency.

The MMI scale was utilised for alert generation, with MMI level 5 selected as an illustrative threshold to demonstrate the alerting process. However, establishing a definitive threshold needs further research for both societal impacts and technical capabilities, underscoring the need for a comprehensive discussion on this topic.

Future Work

In the next phase of our research, we will concentrate on a sequence of strategic undertakings designed to strengthen the capabilities of an operational node-level processing EEWS:

1. **Pilot Implementation and Performance Assessment:** The proposed concept will be piloted by integrating it into several RS4D sensors. We will focus on investigating the PLUM algorithm's performance at the node level, with particular attention to the accuracy of intensity estimations. Data collected from these sensors will be stored in a dedicated database for subsequent analysis to evaluate the system's performance metrics thoroughly.
2. **Network-Wide Implementation:** Upon completing the initial tests and refinements, the algorithm will be rolled out across the entirety of the operational network (45 sensors in the Wellington Region). This crucial step will mark the transition from a controlled testing environment to a real-world application, setting the stage for comprehensive system evaluation.
3. **Alert generation and threshold calibration:** Research efforts will be directed toward fine-tuning the parameters that trigger alert generation. This task will entail a deep dive into alerts' societal impact and the technical nuances associated with threshold calibration, tailored to NZ's unique seismic profile.

4. **Communication Technology Exploration:** Pursuing alternative communication methodologies, such as LoRa, is essential, particularly as an alternative strategy to ensure network operability during traditional communication infrastructure failures.

The conceptual framework detailed in this work-in-progress paper will undergo operational testing, refinement, and expansion. Future results from the experimental network are planned for journal publication to validate the system's effectiveness and provide guidance for similar deployments in seismically active areas.

CONCLUSION

This work represents a notable contribution to EEWSs by integrating node-level processing with the PLUM algorithm. The system's design, emphasising redundancy, cost-efficiency, and reduced latency, addresses the critical need for resilient seismic warning capabilities. Furthermore, the node-level approach emerges as a viable solution following significant earthquakes, where standard communication networks may fail. The potential for sensors to utilise alternative communication methods, such as LoRa, highlights a forward-looking strategy for ensuring continuous alert delivery during aftershocks, enhancing community safety and preparedness. In conclusion, our endeavour to develop a community-engaged node-level EEWS in NZ signifies a step forward in seismic risk management.

FUNDING

This research is funded by the Resilience to Nature's Challenges, QuakeCoRE, a New Zealand Tertiary Education Commission-funded Centre: Publication No. 0954, and by the Toka Tū Ake EQC, New Zealand.

REFERENCES

Allen, R. M., & Melgar, D. (2019). Earthquake early warning: Advances, scientific challenges, and societal needs. *Annual Review of Earth and Planetary Sciences*, *47*, 361–388. https://doi.org/10.1146/annurev-earth-053018-060457

Anthony, R. E., Ringler, A. T., Wilson, D. C., & Wolin, E. (2019). Do low-cost seismographs perform well enough for your network? An overview of laboratory tests and field observations of the OSOP raspberry shake 4D. In *Seismological Research Letters* (Vol. 90, Issue 1). https://doi.org/10.1785/0220180251

Becker, J. S., Potter, S. H., Prasanna, R., Tan, M. L., Payne, B. A., Holden, C., Horspool, N., Smith, R., & Johnston, D. M. (2020). Scoping the potential for earthquake early warning in Aotearoa New Zealand: A sectoral analysis of perceived benefits and challenges. *International Journal of Disaster Risk Reduction*, *51*. https://doi.org/10.1016/j.ijdrr.2020.101765

Brooks, B. A., Protti, M., Ericksen, T., Bunn, J., Vega, F., Cochran, E. S., Duncan, C., Avery, J., Minson, S. E., & Chaves, E. (2021). Robust earthquake early warning at a fraction of the cost: ASTUTI Costa Rica. *AGU Advances*, *2*(3), e2021AV000407.

Chandrakumar, C., Prasanna, R., Stephens, M., & Tan, M. L. (2022). Earthquake early warning systems based on low-cost ground motion sensors: A systematic literature review. *Frontiers in Sensors*, *3*. https://doi.org/10.3389/fsens.2022.1020202

Chandrakumar, C., Tan, M. L., Holden, C., Stephens, M. T., & Prasanna, R. (2023). Performance analysis of P-wave detection algorithms for a community-engaged earthquake early warning system–a case study of the 2022 M5.8 Cook Strait earthquake. *New Zealand Journal of Geology and Geophysics*. https://doi.org/10.1080/00288306.2023.2284276

Chen, D. Y., Hsiao, N. C., & Wu, Y. M. (2015). The earthworm based earthquake alarm reporting system in Taiwan. *Bulletin of the Seismological Society of America*, *105*(2), 568–579. https://doi.org/10.1785/0120140147

Clinton, J., Zollo, A., Marmureanu, A., Zulfikar, C., & Parolai, S. (2016). State-of-the art and future of earthquake early warning in the European region. *Bulletin of Earthquake Engineering*, *14*(9). https://doi.org/10.1007/s10518-016-9922-7

Cremen, G., & Galasso, C. (2020). Earthquake early warning: Recent advances and perspectives. In *Earth-Science Reviews* (Vol. 205). Elsevier B.V. https://doi.org/10.1016/j.earscirev.2020.103184

Dowrick, D. J. (1996). The Modified Mercalli earthquake intensity scale-revisions arising from recent studies of New Zealand earthquakes. *Bulletin of the New Zealand National Society for Earthquake Engineering*, *29*(2).

Fischer, J., Redlich, J.-P., Zschau, J., Milkereit, C., Picozzi, M., Fleming, K., Brumbulli, M., Lichtblau, B., & Eveslage, I. (2012). A wireless mesh sensing network for early warning. *Journal of Network and Computer Applications*, *35*(2), 538–547. http://10.0.3.248/j.jnca.2011.07.016

Fujinawa, Y., & Noda, Y. (2013). Japan's earthquake early warning system on 11 March 2011: Performance, shortcomings, and changes. In *Earthquake Spectra* (Vol. 29, Issue SUPPL.1). https://doi.org/10.1193/1.4000127

GeoNet. (2017). *GeoNet Home*. https://www.geonet.org.nz/

GeoNet, N. Z. (n.d.). *New Zealand Modified Mercalli Intensity Scale*. Retrieved November 10, 2023, from https://www.geonet.org.nz/earthquake/mmi

Holland, A. (2003). Earthquake data recorded by the MEMS accelerometer: Field testing in Idaho. *Seismological Research Letters*, *74*(1). https://doi.org/10.1785/gssrl.74.1.20

Hoshiba, M. (2013). Real-time prediction of ground motion by Kirchhoff-Fresnel boundary integral equation method: Extended front detection method for Earthquake Early Warning. *J. Geophys. Res. Solid Earth*, *118*, 1038–1050. https://doi.org/10.1002/jgrb.50119

Hoshiba, M. (2021). Real-Time Prediction of Impending Ground Shaking: Review of Wavefield-Based (Ground-Motion-Based) Method for Earthquake Early Warning. In *Frontiers in Earth Science* (Vol. 9). Frontiers Media S.A. https://doi.org/10.3389/feart.2021.722784

Hoshiba, M., Iwakiri, K., Hayashimoto, N., Shimoyama, T., Hirano, K., Yamada, Y., Ishigaki, Y., & Kikuta, H. (2011). Outline of the 2011 off the pacific coast of tohoku earthquake (M w 9.0) -earthquake early warning

and observed seismic intensity. *Earth, Planets and Space*, *63*(7), 547–551. https://doi.org/10.5047/eps.2011.05.031

Kodera, Y., Saitou, J., Hayashimoto, N., Adachi, S., Morimoto, M., Nishimae, Y., & Hoshiba, M. (2016). Earthquake early warning for the 2016 Kumamoto earthquake: Performance evaluation of the current system and the next-generation methods of the Japan Meteorological Agency 2016 Kumamoto earthquake sequence and its impact on earthquake science and hazard assessment Manabu Hashimoto, Martha Savage, Takuya Nishimura and Haruo Horikawa 4.Seismology. *Earth, Planets and Space*, *68*(1). https://doi.org/10.1186/s40623-016-0567-1

Kodera, Y., Yamada, Y., Hirano, K., Tamaribuchi, K., Adachi, S., Hayashimoto, N., Morimoto, M., Nakamura, M., & Hoshiba, M. (2018). The propagation of local undamped motion (PLUM) method: A simple and robust seismic wavefield estimation approach for earthquake early warning. *Bulletin of the Seismological Society of America*, *108*(2), 983–1003. https://doi.org/10.1785/0120170085

Kuyuk, H. S., Allen, R. M., Brown, H., Hellweg, M., Henson, I., & Neuhauser, D. (2014). Designing a network-based earthquake early warning algorithm for California: ElarmS-2. *Bulletin of the Seismological Society of America*, *104*(1), 162–173. https://doi.org/10.1785/0120130146

McBride, S. K., Smith, H., Morgoch, M., Sumy, D., Jenkins, M., Peek, L., Bostrom, A., Baldwin, D., Reddy, E., De Groot, R., Becker, J., Johnston, D., & Wood, M. (2022). Evidence-based guidelines for protective actions and earthquake early warning systems. In *Geophysics* (Vol. 87, Issue 1). https://doi.org/10.1190/geo2021-0222.1

Meier, M. A., Heaton, T., & Clinton, J. (2015). The Gutenberg algorithm: Evolutionary Bayesian magnitude estimates for earthquake early warning with a filter bank. *Bulletin of the Seismological Society of America*, *105*(5). https://doi.org/10.1785/0120150098

Moratalla, J. M., Goded, T., Rhoades, D. A., Canessa, S., & Gerstenberger, M. C. (2020). New ground motion to intensity conversion equations (GMICEs) for New Zealand. *Seismological Research Letters*, *92*(1). https://doi.org/10.1785/0220200156

Nakayachi, K., Becker, J. S., Potter, S. H., & Dixon, M. (2019). Residents' Reactions to Earthquake Early Warnings in Japan. *Risk Analysis*, *39*(8), 1723–1740. https://doi.org/10.1111/risa.13306

Noda, S., Yamamoto, S., & Ellsworth, W. L. (2016). Rapid estimation of earthquake magnitude from the arrival time of the peak high-frequency amplitude. *Bulletin of the Seismological Society of America*, *106*(1). https://doi.org/10.1785/0120150108

Prasanna, R., Chandrakumar, C., Nandana, R., Holden, C., Punchihewa, A., Becker, J. S., Jeong, S., Liyanage, N., Ravishan, D., Sampath, R., & Tan, M. L. (2022). "Saving Precious Seconds"—A Novel Approach to Implementing a Low-Cost Earthquake Early Warning System with Node-Level Detection and Alert Generation. *Informatics*, *9*(1), 25. https://doi.org/10.3390/informatics9010025

Santamaria, A. F., Raimondo, P., Tropea, M., De Rango, F., & Aiello, C. (2019). An IoT surveillance system based on a decentralised architecture. *Sensors (Switzerland)*, *19*(6). https://doi.org/10.3390/s19061469

Satriano, C., Elia, L., Martino, C., Lancieri, M., Zollo, A., & Iannaccone, G. (2011). PRESTo, the earthquake early warning system for Southern Italy: Concepts, capabilities and future perspectives. *Soil Dynamics and Earthquake Engineering*, *31*(2), 137–153. https://doi.org/10.1016/j.soildyn.2010.06.008

Suárez, G., Novelo, D., & Mansilla, E. (2009). Performance evaluation of the seismic alert system (SAS) in Mexico City: A seismological and a social perspective. *Seismological Research Letters*, *80*(5). https://doi.org/10.1785/gssrl.80.5.707

Tan, M. L., Brown, A., Stock, K., Becker, J. S., Kenney, C., Lambie, E., Cui, A., & Prasanna, R. (2023). 'Balancing human needs with technology' 1 —a design-led approach for exploring an earthquake early warning system in Aotearoa New Zealand. In *Design for Emergency Management* (pp. 124–140). Routledge. https://doi.org/10.4324/9781003306771-9

Tan, M. L., Prasanna, R., Becker, J. S., Brown, A., Lambie, E., Johnston, D. M., Stock, K., & De Alwis, D. (2021). Outlook for earthquake early warning for Aotearoa New Zealand: Insights from initiating a community-of-practice. *2021 Technical Conference for the New Zealand Society for Earthquake Engineering*, 1–8.

Xi, Z. (2020). *The comparison of decentralized and centralized structure of network communication in different application fields*.

A Study on The Frequency Characteristics of Typhoon Landing in Guangdong, China, Based on Machine Learning Methods

Chenning Pan
Joint International Research Laboratory of Catastrophe Simulation and Systemic Risk Governance, Beijing Normal University, Zhuhai 519087, China;
School of National Safety and Emergency Management, Beijing Normal University, Zhuhai 519087, China
202321099015@mail.bnu.edu.cn

Xiaoyong Ni*
Joint International Research Laboratory of Catastrophe Simulation and Systemic Risk Governance, Beijing Normal University, Zhuhai 519087, China;
School of National Safety and Emergency Management, Beijing Normal University, Zhuhai 519087, China
nixiaoyong@bnu.edu.cn

Ruoxi Lai
Faculty of Arts and Sciences, Beijing Normal University, Zhuhai 519087, China
202211079136@mail.bnu.edu.cn

Dafeng Ma
College of Education for the Future, Beijing Normal University, Zhuhai 519087, China
202211038583@mail.bnu.edu.cn

Xifan Shen
Bay Area Internatinoal Business School,
Beijing Normal University, Zhuhai 519087, China
202111059036@mail.bnu.edu.cn

*Correspondence: nixiaoyong@bnu.edu.cn

ABSTRACT

Located in the southern coastal region of China, Guangdong Province is perennially threatened by typhoons originating from the Northwest Pacific Ocean. Compared to studies focusing on typhoon intensity, paths, and catastrophic effects, research on typhoon landing frequency features remains relatively limited. This study systematically constructs a dataset based on the landing typhoon frequency of Guangdong over 71 years (as the target variable) and a set of 88 atmospheric circulation indices as well as 26 sea surface temperature indices (as the potential feature variables). Two machine learning models, including Random Forest Regression (RR) and Support Vector Regression (SVR), are used to predict the frequency of typhoons making landfall in Guangdong, and the results show that both models have good characterisation ability while the RR model has slightly better fitting performance on the training set. This study provides scientific support for understanding the characteristics of typhoons in Guangdong and better responding to typhoon disasters.

Keywords

Tropical cyclone, frequency characteristics, machine learning, vulnerability functions.

INTRODUCTION

According to the definition of the World Meteorological Organization, a tropical cyclone with sustained wind speeds above level 12 (i.e. 32.7 to 41.4 meters per second) at its center is called a typhoon or hurricane, where generally the former is used to refer to tropical cyclones generated in the Pacific Ocean, while the latter is used to refer to those generated in the Atlantic Ocean. According to the traditional Chinese custom of naming tropical cyclones, tropical cyclones are collectively referred to as typhoons in the following. The 2022 Global Natural Disaster Assessment Report points out that there are a total of 66 major storm disasters (typhoons and hurricanes) worldwide in 2022, accounting for approximately 21% of the total frequency of major disasters; 16.93 million people are affected, resulting in direct economic losses of approximately 131 billion US dollars (Academy of Disaster Reduction and Emergency Management et al., 2023). The impact and challenges caused by typhoon disasters cannot be ignored and are a global problem facing humanity. China is located in the northwestern part of the Pacific Ocean, which is a high risk area for typhoon disasters. In the 71 years from 1949 to 2019, there are 2,333 typhoons in the Northwest Pacific, of which 636 have landed in China, accounting for 27.3% of the total (Liu et al., 2012; Su et al., 2021). Guangdong Province is located in the coastal area of southern China and is the main region where typhoon makes landfall in China, making it one of the provinces most severely affected by typhoons (Zhang et al., 2009; Chou et al., 2018). In 2018, Typhoon Mangkhut makes landfall in Guangdong with 14 gusts of wind, affecting 3.066 million people, causing the collapse of 200,000 houses and direct economic losses of 13.29 billion yuan. In early September 2023, Typhoon Sura and Haikui make landfall in Guangdong, causing record-breaking rainfall in various parts of the province. The disasters have caused serious consequences, including road collapses, interruptions in communication and power supply, house collapses, fallen trees, and damage to water conservancy projects, particularly in Shenzhen and Yangjiang.

With the severity of global climate change, the increasing intensity of worldwide natural disasters such as typhoons is causing progressively greater losses globally. Typhoon characteristics prediction has become a major focus of research. In general, the studies on typhoon characteristics prediction include typhoon frequency prediction (Jia et al., 2014; Hai et al., 2019), track prediction (Murakami et al., 2010; Hsan et al., 2021), landfall prediction (Wahiduzzaman et al., 2017; Kumar et al., 2021), wind field strength prediction (Kotal et al., 2020; Chen et al., 2023), rain field strength prediction (Matyas, 2010), and so on. The study of the characteristics of typhoon landing frequency can understand the law and mechanism of typhoon activities and the impact of climate change on typhoon activities, better early warning and risk management, so this paper mainly focuses on the prediction of typhoon landing frequency.

The traditional methods for studying the characteristics of typhoon landing frequency mainly include two categories: statistical analysis methods and numerical analysis methods. Statistical methods do not focus on physical principles but on the exploration of data patterns in historical disasters. Yang et al. find that the landfall typhoons numbers as well as the south-north range of landfall typhoons have decreased by using the data from the Tropical Cyclone Yearbooks (Yang et al., 2009). Numerical analysis methods are built based on principles of atmospheric dynamics and thermodynamics to discretize and simulate the atmosphere for weather forecasting. Nie et al. find that the average intensity and number of typhoons landing in China under three RCPs will increase in varying degrees in the future, based on outputs from the global climate model HadGEM2-ES in CMIP5 (Nie et al., 2023). With the popularization of artificial intelligence technology, machine learning has gradually become a popular method in typhoon characteristics prediction research due to their ability to process large amounts of complex meteorological data and capture their nonlinear and non parametric relationships. Richman et al. implement a Support Vector Regression model based on the May-September Nino index, a novel El Nino, a quasi-biennial oscillation in the stratosphere, and a Southern Hemisphere ring mode to estimate the frequency of typhoons in the Australian typhoon season from December to April (Richman et al., 2012). Tan et al. select environmental fields such as sea surface temperature, sea level pressure, Nino 3.4 index, wind shear, vorticity, subtropical high pressure and sea ice cover, and establish a random forest model based on the prediction of tropical cyclone frequency in the Northwest Pacific Ocean (Tan et al., 2018). Chen chooses six types of forecast elements, including temperature field, altitude field, sea-level pressure, specific humidity field, meridional wind field vertical shear, and sea surface temperature field to perform single machine learning model predictions and integrated predictions (Chen, 2022).

Overall, compared to traditional weather prediction methods such as numerical models, machine learning methods are still in the exploratory stage of typhoon characteristics prediction research. Its application research scope is wide, but it mainly focuses on the study of typhoon intensity, weather and its catastrophic effects such as wind speed and rain intensity (Chen et al., 2020). There is still relatively little research on typhoon landing frequency. Studying the frequency of landfall can help understand the characteristics and trends of typhoon frequency in different regions, help governments and communities better prepare and respond to possible typhoon disasters, conduct more effective meteorological disaster risk assessments, and reduce the impact of disasters on people's

lives and property. Therefore, studying the frequency of typhoons is of great significance for guiding regional disaster prevention and reduction.

THE OVERALL STRUCTURE

The general structure of this study is shown in Figure 1. Guangdong Province has been selected as the study area, which is one of the most typhoon-prone provinces in mainland China. Typhoon track data from 1952 to 2022 and a potential set of feature variables have been collected. Through data processing, datasets of both feature variables and target variables have been obtained, used for training two machine learning models. The ability of the two models to characterise and predict the frequency of typhoons making landfall in Guangdong Province is compared. The results of this paper help to understand the characteristics of typhoons in Guangdong Province and provide reference and technical support for subsequent studies on typhoon landing frequency predicion, as well as help to improve the government's capabilities in response to typhoons.

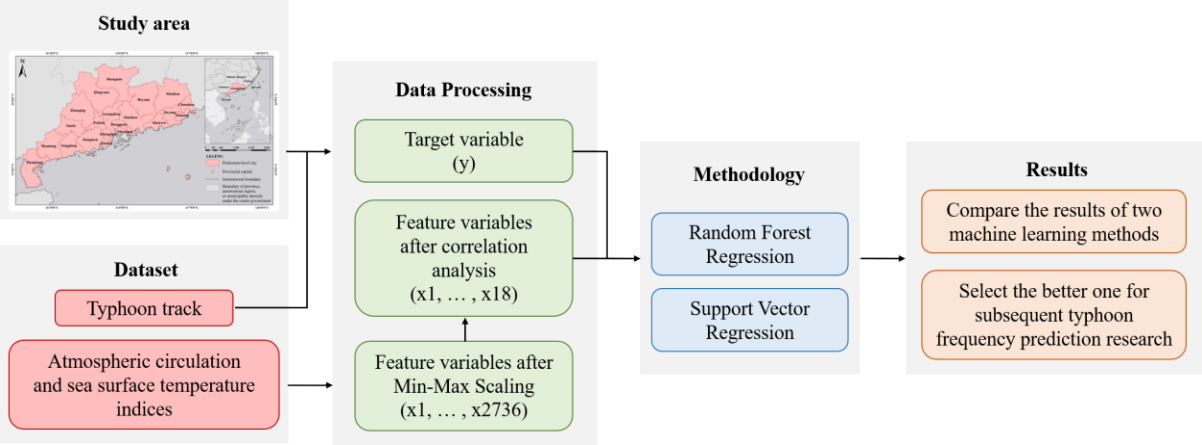


Figure 1. The Overall Structure of This Study

STUDY AREA AND DATASETS

Study Area

Guangdong Province is a provincial-level administrative region in China. It is located along the coast of the South China Sea, bordered by Hong Kong, Macau, Guangxi, Hunan, Jiangxi and Fujian, and across the sea from Hainan Province, with a total land area of 179,800 square kilometers. As of October 2022, Guangdong is composed of 21 prefecture-level cities, 65 districts, 20 county-level cities, 34 counties, and 3 autonomous counties. In 2022, the gross domestic product (GDP) of Guangdong Province reached 12,911.86 billion yuan, with a per capita GDP of 101,905 yuan. At the end of 2022, the permanent population of Guangdong was 126.568 million, with an urbanization rate of 74.79%. Figure 2 illustrates the study area in this paper.

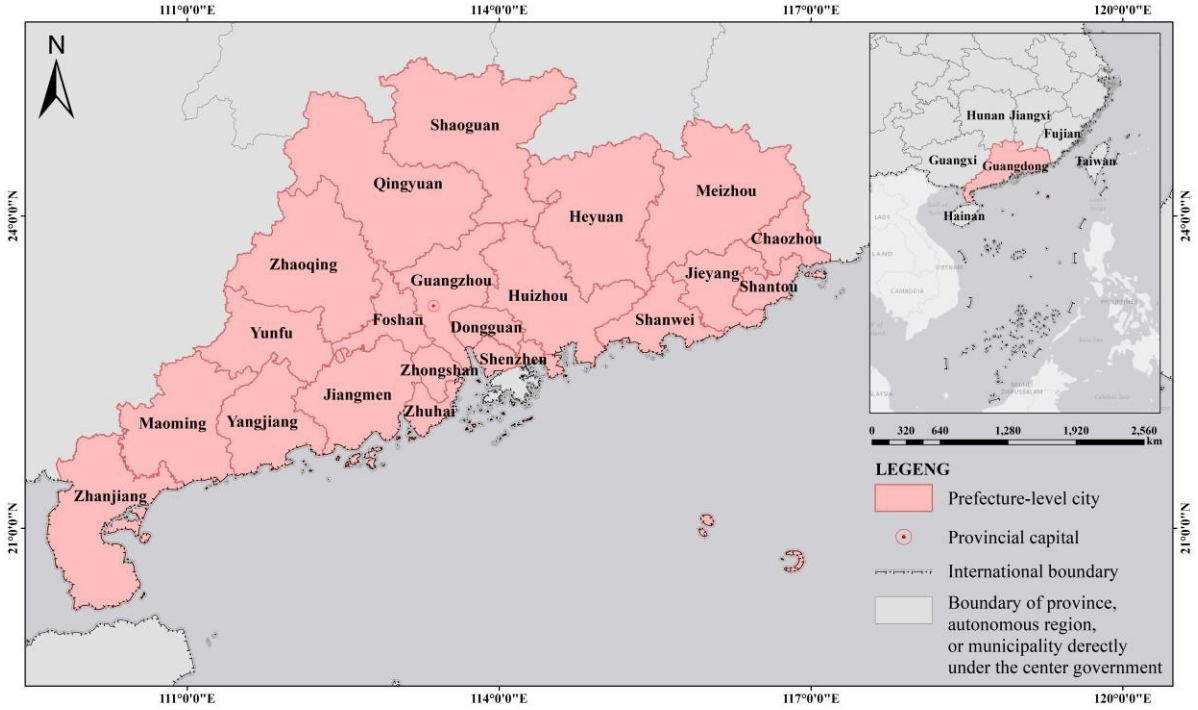


Figure 2. The Study Area in This Study

Data Sources

The typhoon track data is sourced from the Best Track Data provided by the China Meteorological Administration Tropical Cyclone Data Center (CMA, 2023; Ying et al., 2014; Lu et al., 2021). This dataset provides basic information on the position and intensity of tropical cyclones in the northwest Pacific (including the South China Sea, north of the equator, and west of 180°E longitude) every 6 hours since 1949. Starting from 2017, for typhoons that make landfall in China, the time frequency of the best track data has been increased to once every 3 hours within the 24-hour period before their landfall. Since 2018, for typhoons making landfall in China, the time frequency of the best track data is increased to every 3 hours within the 24 hours before their landfall and during their land activities In China.

Regarding the potential feature factors influencing typhoon frequency, the monthly data of 88 atmospheric circulation indices and 26 sea surface temperature indices provided by the National Climate Center from 1951 to 2022 are considered by this study, mainly based on the research of Rong et al. (National Climate Center, 2023; Rong et al., 2023).

Raw Data Processing

Feature Variables Processing

For the feature variables, this study utilizes monthly atmospheric circulation indices and sea surface temperature indices for the years 1951-2022. The entire feature variable dataset comprises 71 entries (an entry per year from 1952 to 2022), with each entry including 88 atmospheric circulation indices such as the Northern Hemisphere Subtropical High Area Index, as well as 26 sea surface temperature indices like the NINO 1+2 SSTA Index, for a total of 24 months corresponding to the target variable's current year and previous year, resulting in a total of 2736 [=(88+26)*24] feature variables per entry.

In the feature variable dataset, the differences in meanings of various variables and units of measurement can affect the performance of machine learning algorithms. Therefore, it is necessary to normalize data from different dimensions, allowing for comparison and interpretation on the same scale. In this study, Min-Max Scaling is employed to normalize the feature variable dataset.

The Min-Max Scaling method scales the original data proportionally to fit within the range [0, 1]. The specific formula for this process is as follows:

$$x_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (1)$$

Here, x_{ij} denotes the value of the j-th variable for the i-th year, $\min(x_j)$ and $\max(x_j)$ are the minimum and maximum values of the j-th variable in the whole dataset, and x'_{ij} denotes the normalized value after processing.

Acquisition and Analysis of the Target Variable

Based on the best track data of tropical cyclones, which provides latitude and longitude information at regular intervals for each typhoon, combined with the longitude and latitude range of Guangdong Province, whether the typhoon made landfall in mainland China from Guangdong Province is determined. Here, landing in Guangdong refers to the first landfall location of a typhoon in China being Guangdong, the situation where a typhoon lands in other provinces of China first and then makes a secondary landfall in Guangdong is not considered. Summarizing the typhoons landing in Guangdong each year, the frequency data of typhoons making landfall in Guangdong from 1952 to 2022 (target variable) is obtained. It reveals that a total of 250 typhoons made landfall in Guangdong during this period, with an average of 3.52 typhoons per year and a standard deviation of 1.71. The highest annual typhoon frequency reached seven, while the lowest was one. Figure 3 reflects the annual changes in the frequency of typhoons landing in Guangdong Province. The linear trendline indicates a significant downward trend in the number of typhoons making landfall in Guangdong Province over the 71 years. The 5-year moving average curve shows that the landing typhoon frequency exhibits overall fluctuating characteristics, entering a period of reduced frequency starting from the late 20th century.

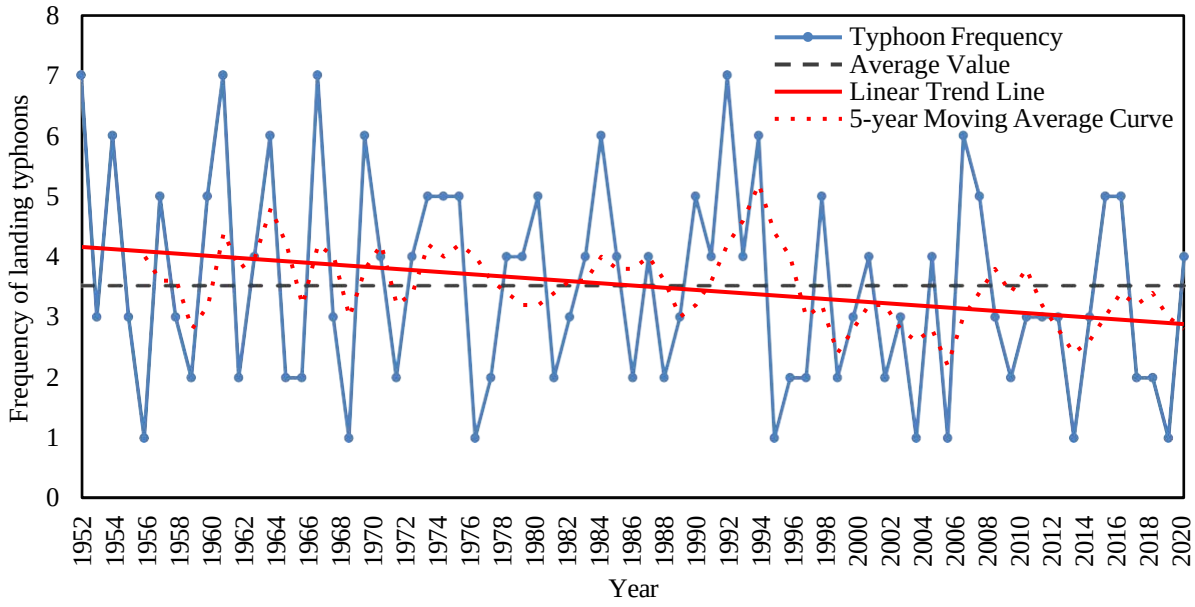


Figure 3. The Annual Distribution of Landing Typhoon Frequency in Guangdong Province

METHODOLOGY

Random Forest Regression

Random Forest Regression (RR) is an important application branch of Random Forest. The RR model builds multiple unrelated decision trees by randomly extracting samples and features, following the principles of Bagging, and obtains predictions in parallel. Each decision tree can produce a prediction result from the extracted samples and features, and the regression prediction result of the whole forest is obtained by averaging the results of all trees. This model improves the Bagging ensemble algorithm by incorporating random attribute selection alongside the bootstrap sampling method during the training of decision trees. This introduces diversity among weak learners, enhancing the model's generalization performance.

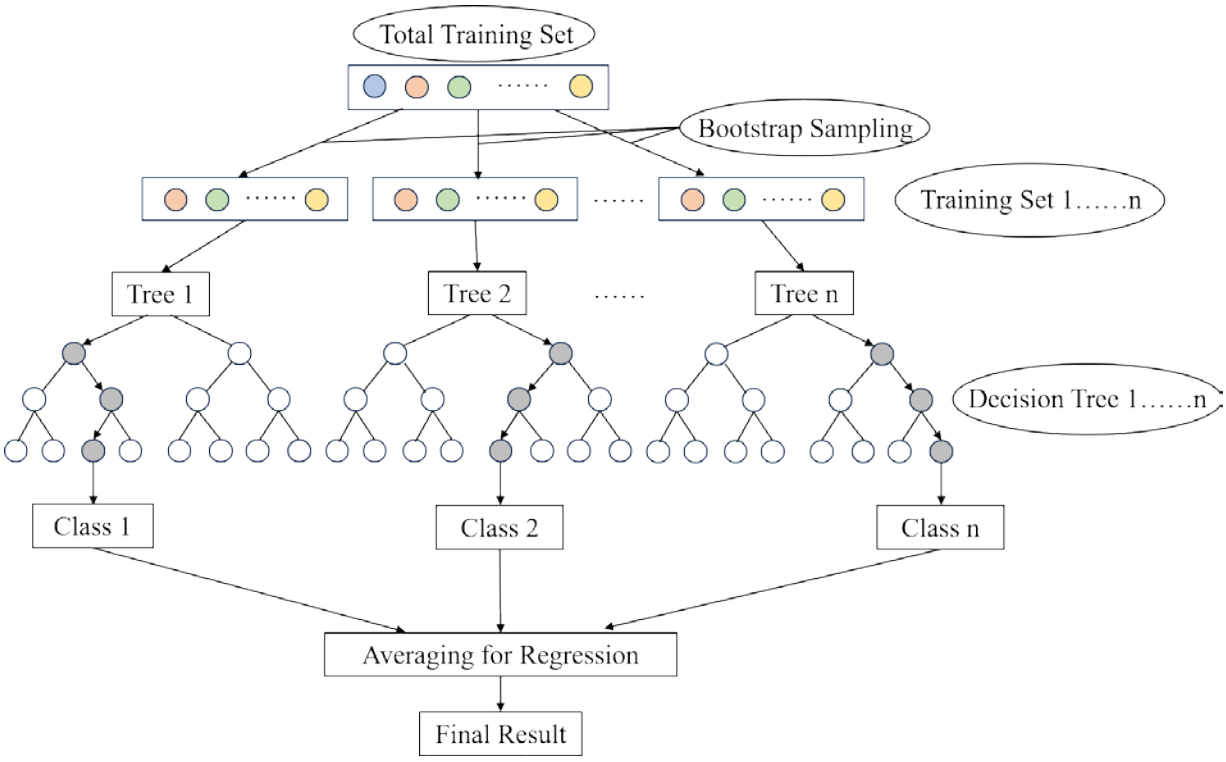


Figure 4. Flowchart of Random Forest Regression Algorithm

Support Vector Regression

Support Vector Regression (SVR) is a supervised learning method that uses Support Vector Machines (SVM) for regression analysis. The goal of SVR is to find the optimal hyperplane, allowing for a certain amount of error, such that the sample points are as close to this hyperplane as possible. SVR achieves this by optimizing the following convex quadratic programming problem:

$$\min_{\omega,b,\xi,\xi} \frac{1}{2}||\omega||^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i) \tag{2}$$

Here, ω denotes the weight vector of the hyperplane, b denotes the intercept, ξ and ξ denote slack variables allowing some data points to be within the margin or error range, respectively, and C denotes the regularization parameter controlling the trade-off between margin size and error.

The objective of SVR is to minimize the loss function mentioned above while satisfying the following constraints:

$$yi - \langle \omega, \phi(x_i) \rangle - b \leq \epsilon + \xi_i \tag{3}$$

$$\langle \omega, \phi(x_i) \rangle + b - yi \leq \epsilon + \xi_i \tag{4}$$

$$\xi_i, \xi_i \geq 0 \tag{5}$$

These constraints ensure that data points are within the margin or error range. The parameter ϵ controls the size of the margin. SVR finds the optimal hyperplane by minimizing the loss function, enabling it to make predictions.

Metrics for Evaluating Machine Learning Performance

In order to compare the learning effects of the two machine learning models, Mean Squared Error (MSE), Root Mean Squard Error (RMSE) and Mean Absolute Error (MAE) are chosen as the evaluation metrics for the training results. Smaller MSE, MAE and RMSE suggest better fit of the model to the data. The formulas to calculate MSE, RMSE and MAE are as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{8}$$

In which, n denotes the number of samples, y_i denotes the actual values, and $\hat{y}_i$ denotes the predicted values of the model.

RESULTS

Correlation Analysis

To prevent overfitting and reduce redundant information caused by numerous feature variables, a correlation analysis between the target variable and feature variables is performed, yielding correlation coefficients of each feature variable to the target variable. Based on the correlation coefficients, features with a strong correlation with the target variable have higher priority to be selected thereby simplifying the feature variables dataset, which is recommended by Rong et al. as a simple but effective way of parameter selection (Rong et al., 2023). Given a dataset D = $\{(X_K, Y_K)\}$, where $K \in n$, the correlation coefficient r is defined as:

$$r = \frac{\sum_{K=1}^{n}(X_K - \bar{X})(Y_K - \bar{Y})}{\sqrt{\sum_{K=1}^{n}(X_K - \bar{X})^2}\sqrt{\sum_{K=1}^{n}(Y_K - \bar{Y})^2}} \tag{9}$$

where $\bar{X}$ denotes the mean of X and $\bar{Y}$ denotes the mean of Y.

After calculating the correlation coefficients for each feature variable to the target variable, those with missing values (originally marked as -999 in the raw data) are removed. The remaining features are sorted in descending order based on the absolute values of their correlation coefficients. Subsequently, multiple tests are conducted based on the changing dataset, where every three sequential cumulative increases in the number of feature variables, and the dataset composed of the first 18 feature variables is discovered to have the better prediction effect. Therefore, in this paper, 18 feature variables are set as shown in Table 1, and due to space limitations, the process of selecting the number of feature variables is omitted. These 18 feature variables respectively represent the sea surface temperature indices (x2723, x2734, x1902, x2620, x2724, x1580, x2609, x2718, x2610, x2604, x2506, x1694, x2512) and atmospheric pressure indices (x2575, x2573, x2337, x411, x1605) related to typhoon formation (For detailed explanations of these indices, please refer to the National Climate Center's website). Furthermore, only x411 is a feature variable from the previous year, affecting this year's frequency of typhoon landing. Therefore, it can be observed that sea surface temperature factors have a greater impact on the frequency of typhoon landing in Guangdong, and the index of the same year has more influence compared to the index of the previous year. The dataset of feature variables together with the target variable will serve as the initial dataset for machine learning models. The dataset is split into 80% for training and 20% for testing.

Table 1. The Top Eighteen Feature Variables Ranked by the Absolute Values of Their Correlation Coefficients

Number	Feature variable identifier	Feature variable name	Absolute value of the correlation coefficient
1	x2723	Indian Ocean Warm Pool Area Index in December of the same year	0.4015
2	x2734	Indian Ocean Basin-Wide Index in December of the same year	0.3995
3	x1902	Scandinavia Pattern , SCA in May of the same year	0.3949
4	x2620	Indian Ocean Basin-Wide Index in November of the same year	0.3766
5	x2724	Indian Ocean Warm Pool Strength Index in December of the same year	0.3731

6	x1580	Tropical Northern Atlantic SST Index in February of the same year	0.3659
7	x2609	Indian Ocean Warm Pool Area Index in November of the same year	0.3605
8	x2575	India-Burma Trough Intensity Index in November of the same year	0.3564
9	x2718	NINO B SSTA Index in December of the same year	0.3514
10	x2610	Indian Ocean Warm Pool Strength Index in Novermber of the same year	0.3483
11	x2604	NINO B SSTA Index in Novermber of the same year	0.3385
12	x2573	Tibet Plateau Region 1 Index in Novermber of the same year	0.3376
13	x2337	Northern Hemisphere Polar Vortex Central Latitude Index in September of the same year	0.3373
14	x2506	Indian Ocean Basin-Wide Index in October of the same year	0.3373
15	x411	Antarctic Oscillation, AAO in April of the previous year	0.3334
16	x1605	South China Sea Subtropical High Area Index in March of the same year	0.3332
17	x1694	Tropical Northern Atlantic SST Index in March of the same year	0.3298
18	x2152	Western Hemisphere Warm Pool Index in July of the same year	0.3293

Random Forest Regression Model Prediction

Key parameters affecting the fitting performance of the RR model, namely n_estimators (number of trees), max_depth (maximum depth of the tree), min_samples_split (minimum samples required for a node to split), and min_samples_leaf (minimum samples required for a leaf node), are performed parameter tuning by using a parameter grid which is formed by commonly used numerical values respectively. After multiple rounds of optimization and combining the fitting results, the values 200, 20, 10, and 4 are chosen for the respective parameters.

Due to the random features inherent in the RR model, the results obtained from each training iteration are not entirely consistent. Therefore, Table 2 represents the average training results after 20 repetitions. The model exhibites good performance on the training set, where the prediction deviation for typhoon landing frequency is around 1. However, the performance on the test set is relatively poor, with the values of three metrics significantly higher than on the training set. Comparison of the actual and predicted values for the training and test sets from the 20th training iteration is shown in Figure 5. The RR model demonstrates a consistent trend with the actual values in the training set, showing relatively high concordance with peak and trough values in certain years. On the other hand, the results for the test set, highlighted with a deep red background in the Figure 5, exhibites suboptimal trends and extreme value predictions. The maximum prediction error for the number of landing typhoons reaches 2.4.

Table 2. The Results of RR Training Evaluation Metrics

Type of dataset	Evaluation indicators	Values
Training set	MSE	1.0857 ± 0.0152
	RMSE	1.0419 ± 0.0073

	MAE	0.8758 ± 0.0096
	MSE	1.8042 ± 0.0511
Test set	RMSE	1.3431 ± 0.0192
	MAE	1.1669 ± 0.0178

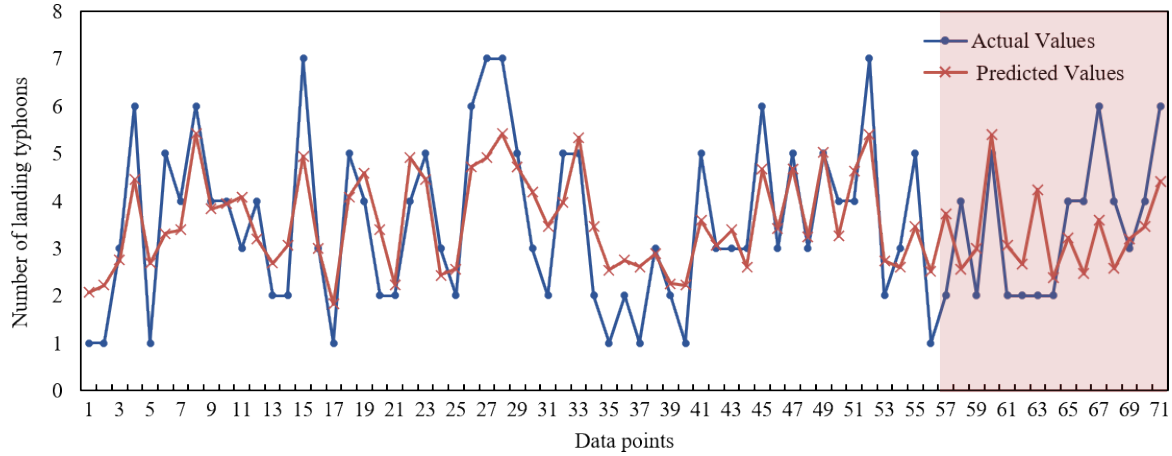


Figure 5. RR Training Results Plot for Training and Test Sets (Highlighted in Deep Red Background for Test Set Training Results)

The average feature importance values from the 20 training iterations are illustrated in Figure 6. Among them, the four feature variables, Scandinavia Pattern, SCA in May of the same year (x1902), Antarctic Oscillation, AAO in April of the previous year (x411), Tropical Northern Atlantic SST Index in February of the same year (x1580), and Tibet Plateau Region 1 Index in Novermber of the same year (x2573), demonstrate higher importance in influencing the overall training performance of the model, with importance values of 0.2097, 0.1159, 0.1154, and 0.0899, respectively. The relevant results show that time coefficients of the ninth mode from empirical orthogonal function analysis (EOF) of the normalised 500hPa height field in the region 20°N-90°N, 0-360°in May as well as regional average of sea surface temperature distance levels in the region 5.5°N-23.5°N, 57.5°W-15°W in February of the same year (x1902, x1580) and normalised series of time coefficients of the first mode from EOF of the anomalous field at 700hPa height in the region 20°-90°S, 0-360°in April of the previous year (x411) will have a relatively important impact on the number of typhoon landing in Guangdong Province in a year.

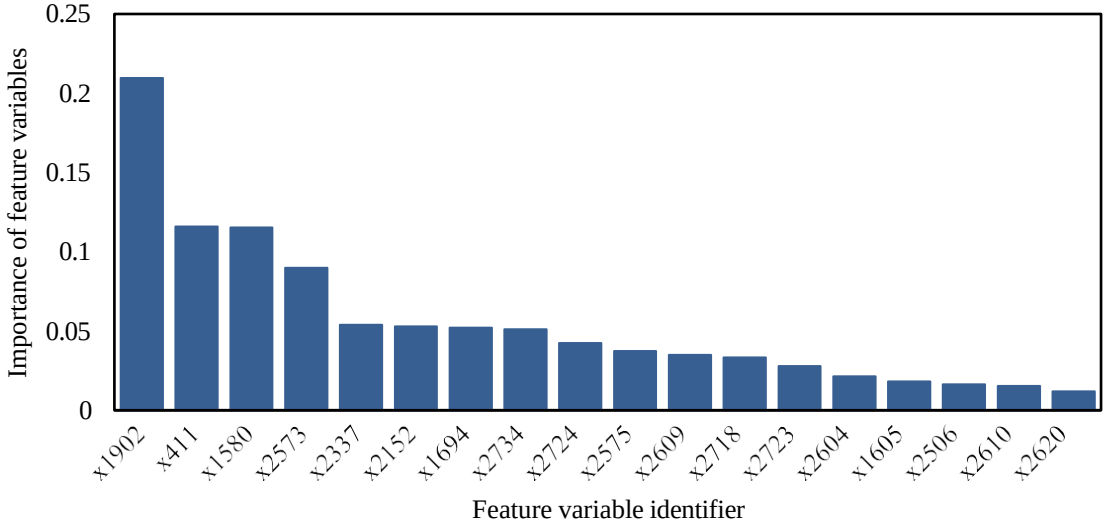


Figure 6. Importance Level of Eighteen Feature Variables

Support Vector Regression Model Prediction

In SVR model, the parameters affecting training performance include kernel (kernel function), C (penalty term), epsilon (tolerance), and gamma (affecting the width of the kernel function). Similar to Random Forest Regression, common values for these parameters are selected to form a parameter grid, picking the best parameter combination through exhaustive methods. SVR does not involve the concept of "random" features like Random Forest, so the optimal parameter set is obtained through one training: C is set to 10, epsilon is set to 1, gamma is set to 0.1, and the kernel is set to sigmoid.

The training results for the training set and test set are shown in Table 3. It can be observed that the training performance of SVR on the training set composed of these 18 feature variables is not very ideal compared to RR, with an average of 1.18 typhoon landing frequency prediction errors according to MAE, which is higher than RR. However, the test set fitting effect is good, with the values of three metrics lower than RR. Combining Figure 7, SVR's prediction values on the training set align well with the overall trend of the actual values, but the fitting performance for the peak and trough values is relatively poor, even showing completely opposite predicted values to the actual values in some years. The test set results highlighted with a deep green background show suboptimal trend and extreme value performance, with a maximum prediction error of 2.29 for the number of landing typhoons.

Table 3. The Results of SVR Training Evaluation Metrics

Type of dataset	Evaluation indicators	Values
Training set	MSE	1.9894
	RMSE	1.4105
	MAE	1.1840
Test set	MSE	1.5054
	RMSE	1.2269
	MAE	1.0259

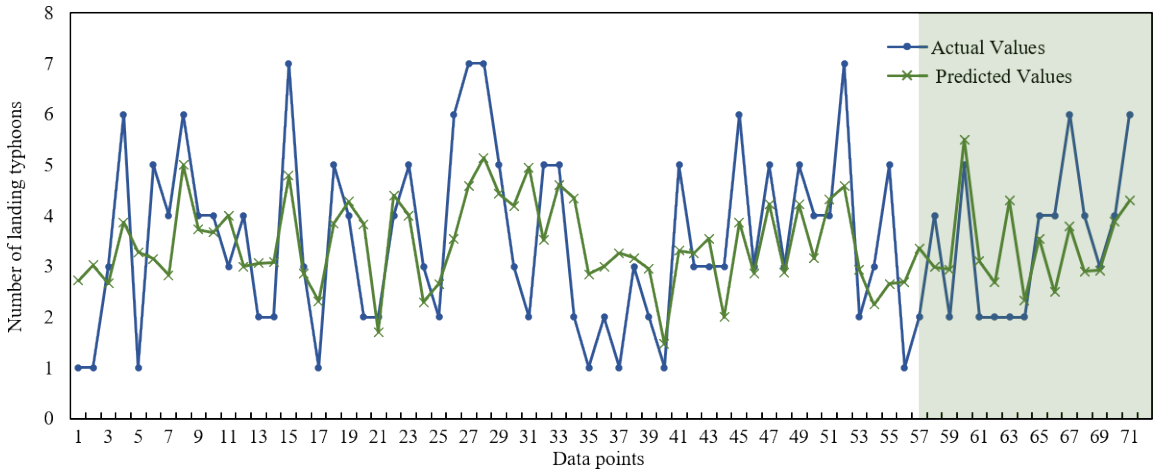


Figure 7. SVR Training Results Plot for Training and Test Sets (Highlighted in Deep Green Background for Test Set Training Results)

Comparison of Model Performance after Simple Secondary Selection

Usually, multicollinearity among feature variables can lead to model instability and excessive sensitivity to input variables. It is necessary to adopt methods to select optimal feature variables from a pre-selected set to enhance model performance and interpretability. Based on the idea of Rong et al., the author's team proposes a new approach: conducting preliminary selection of feature variable dataset by correlation analysis and secondary selection based on importance obtained through the RR method mentioned earlier to simplify the number of feature variables in the dataset, and then performing regression training based on two machine learning models. Table 4 below shows the 8 feature variables with importance values exceeding 0.05, which will constitute a new feature variable dataset for training both models.

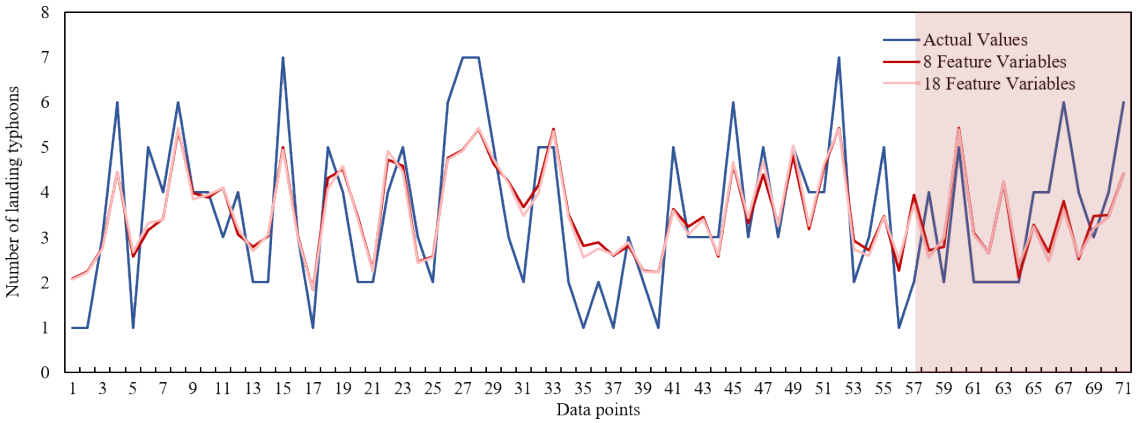
Based on this idea, the author's team does a preliminary exploration The training results of the two models are shown in Table 5. Compared to the dataset with 18 feature variables, the values of MSE, RMSE, and MAE on the test set all decrease, indicating that the dataset with 8 feature variables shows some improvement in the RR model's fitting performance on the test set. However, in the training set, the simplified dataset has slightly worse model fitting performance than the original dataset. For the SVR model, the simplified dataset does not improve its fitting performance. Comparing the actual and predicted values of the two models on the training and test sets with different feature variable datasets, the simplified dataset does not play much role in two model fitting performances.

Table 4. The Eight Feature Variables with Importance Values Exceeding 0.05

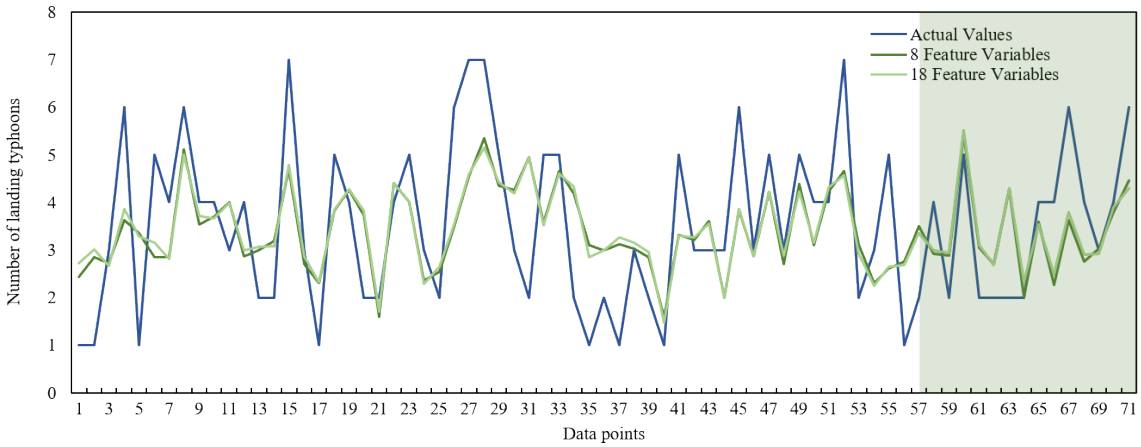
Number	Feature variable identifier	Feature variable name	Importance value of the characteristic indicator
1	x1902	Scandinavia Pattern , SCA in May of the same year	0.2097
2	x411	Antarctic Oscillation, AAO in April of the previous year	0.1159
3	x1580	Tropical Northern Atlantic SST Index in February of the same year	0.1154
4	x2573	Tibet Plateau Region 1 Index in Novermber of the same year	0.0899
5	x2337	Northern Hemisphere Polar Vortex Central Latitude Index in September of the same year	0.0540
6	x2152	Western Hemisphere Warm Pool Index in July of the same year	0.0530
7	x1694	Tropical Northern Atlantic SST Index in March of the same year	0.0521
8	x2734	Indian Ocean Basin-Wide Index in December of the same year	0.0511

Table 5. The Training Evaluation Results of the Two Models

Type of dataset	Evaluation indicators	RR		SVR	
		18 feature variables	8 feature variables	18 feature variables	8 feature variables
Training set	MSE	1.0857 ± 0.0152	1.1033 ± 0.0163	1.9894	2.0137
	RMSE	1.0419 ± 0.0073	1.0504 ± 0.0078	1.4105	1.4191
	MAE	0.8758 ± 0.0096	0.8832 ± 0.0060	1.1840	1.1913
Test set	MSE	1.8042 ± 0.0511	1.7534 ± 0.0382	1.5054	1.5912
	RMSE	1.3431 ± 0.0192	1.3241 ± 0.0144	1.2269	1.2614
	MAE	1.1669 ± 0.0178	1.1426 ± 0.0139	1.0259	1.0301



(a) RR



(b) SVR

Figure 8. The Training Results of the Two Models on the Training Set and Test Set Based on Two Datasets (Highlighted in Background for Test Set Training Results)

The results show that the simplification method described above applied to the feature variable dataset does slightly improve the model fitting performance, but the improvement is limited. This suggests that the combined indicator selection approach of initial screening based on relevance and secondary selection based on importance has some potential, but its utility needs further exploration.

LIMITATIONS AND FUTURE WORK

In the selection of machine learning prediction models, due to time and space constraints, this paper only utilizes common machine learning regression models-Random Forest Regression and Support Vector Regression. However, there are many different regression models in machine learning, and for a more comprehensive and refined application to train and predict the landing typhoon frequency in Guangdong Province, other machine learning methods and even deep learning or large model learning methods can be considered as future research methods. By comparing the training results, the optimal method for predicting the landing typhoon frequency in Guangdong Province may be identified. After continuous optimization of the model, further practice can be considered - translating research findings into an information system. This system would display the predicted typhoon landing frequency for each region based on input atmospheric and sea temperature indices. We briefly conceptualized the initial interface and functionality of the system as shown in the figure below (the image results displayed are not accurate). The development of this visualization system can help for typhoon crisis management.

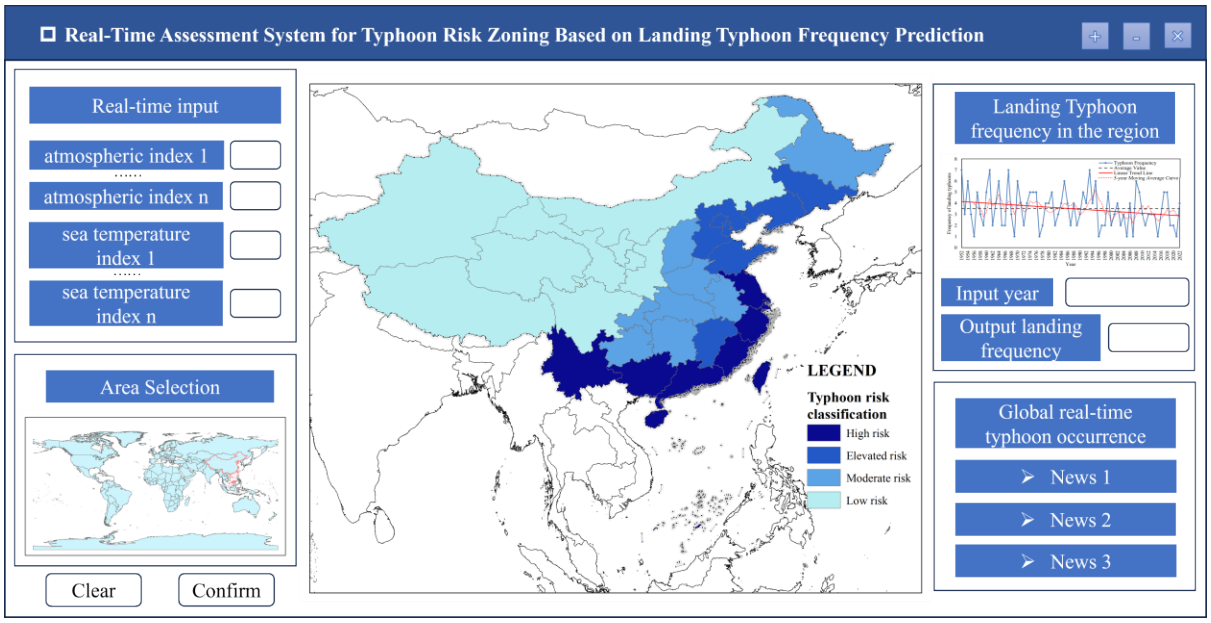


Figure 9. The Initial Concepts for the System Interface

CONCLUSION

Tropical cyclones are one of the destructive and globally impactful natural disasters. With the increasingly severe global climate change, the frequency of extreme weather events is rising, and tropical cyclone disasters have become a significant global concern. This study is based on collected data of the best tracks of typhoons from 1952 to 2022 and monthly atmospheric circulation and sea surface temperature indices from 1951 to 2022, constructing target variable and feature variable datasets separately. These datasets are used for training Random Forest Regression (RR) and Support Vector Regression (SVR) models, allowing a comparison of the fitting performance of the two machine learning models.

The research findings show that the frequency of typhoons making landfall in Guangdong Province has shown a decreasing trend from 1952 to 2022, particularly entering an interdecadal period of scarcity began at the end of the 20th century. Based on the training results of RR and SVR models on the dataset with 18 feature variables, RR demonstrates better fitting performance on the training set compared to SVR, while the opposite is observed on the test set. Both models exhibit good fitting trends on the training set, but RR shows higher sensitivity to peaks and valleys than SVR. On the test set, both models demonstrate relatively less-than-ideal fitting in terms of trends and extreme values. Generally speaking, the values of MSE, RMSE and MAE are small, so it can be seen that the prediction of the number of landing typhoons in this paper is accurate, basically able to predict the exact number of landing typhoon, with a possible error of no more than 2 typhoons. Additionally, analysis of the importance of the feature variables by RR model shows that SCA in May and Tropical Northern Atlantic SST Index in February of the same year and AAO in April of the previous year have a relatively important impact on the number of typhoon landing in Guangdong Province in a year. After the second screening and simplification of the dataset using the RR method, there is not a significant improvement in the fitting performance of the two models. The RR model shows some improvement, but it has the opposite effect on the SVR model, leading to a decrease in performance. Therefore, the combined indicator selection approach of initial screening based on relevance and secondary selection based on importance has some potential, which needs further exploration.

The research process and related results provide some references and guidance for further research and prediction of landing typhoon frequency and other characteristics of typhoons. This paper also helps the government departments to assess the risk of typhoons in the region and to improve the capabilities of emergency response management to disaster risks.

ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (Grant No. 72304038), Guangdong Basic and Applied Basic Research Foundation (Grant No. 2022A1515110049), Young Innovative Talents Project from Department of Education of Guangdong Province (Grant No. 2022KQNCX157).

DATA AVAILABILITY STATEMENT

All the data used in this article are provided with references and links at the manuscript.

WiP Paper – IT Solutions for Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

REFERENCES

Academy of Disaster Reduction and Emergency Management, National Disaster Reduction Centre of China, International Federation of Red Cross and Red Crescent Societies (IFRC), Beijing Normal University. (2023). 2022 Global natural disaster assessment report. https://www.preventionweb.net/publication/2022-global-natural-disaster-assessment-report

Liu, H., & Liu, R. (2012). Tropical Cyclones Activities in the Western North Pacific from 1990 through 2009. Resources Science, *34*(2), 242-247.

Su, H., Yuan, L., Wang, M., Dong, G., & Fei, X. (2021). Characteristics and disaster analysis of tropical cyclones landing in China in 1949-2019. *Journal of Applied Oceanography*, *40*(3), 382-387.

Zhang, L., Liu, M., Quan, R., Lu, M., Wang, J., & Xu, S. (2009). Characteristics and disaster valuation of the tropical cyclones in southeast coastal areas of China. *Journal of East China Normal University (Natural Science)*, *2*, 41-49.

Chou, J. M., Ban, J. H., Dong, W. J., Hu, C., & Dai, R. (2018). Characteristics analysis and assessment of economic damages caused by tropical cyclones in Guangdong Province. *Chinese Journal of Atmospheric Sciences*, *42*(2), 357-366.

Jia, X., Chen, L., & Luo, J. (2014). Climate Prediction Experiment for Tropical Cyclone Genesis Frequency Using the Large-Scale Circulation Forecast by a Coupled Global Circulation Model. *Journal of Tropical Meteorology*, *20*(2), 103–111.

Hai, Y., & Chen, G. H. (2019). Prediction of frequency of tropical cyclones forming over the Western North Pacific using an artificial neural network model. *Clim Environ Res*, *24*(3), 324-332.

Murakami, H., & Wang, B. (2010). Future change of North Atlantic tropical cyclone tracks: Projection by a 20-km-mesh global atmospheric model. *Journal of Climate*, *23*(10), 2699-2721. https://doi.org/10.1175/2010JCLI3338.1

Hsan, T. Z., & Sein, M. M. (2021). Combining Support Vector Machine and Polynomial Regressing to Predict Tropical Cyclone Track. In *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)* (pp. 220-221). IEEE. https://doi.org/10.1109/LifeTech52111.2021.9391780

Wahiduzzaman, M., Oliver, E. C., Wotherspoon, S. J., & Holbrook, N. J. (2017). A climatological model of North Indian Ocean tropical cyclone genesis, tracks and landfall. *Climate Dynamics*, *49*, 2585-2603. https://doi.org/10.1007/s00382-016-3461-4

Kumar, S., Biswas, K., & Pandey, A. K. (2021). Predicting landfall's location and time of a tropical cyclone using reanalysis data. In *Artificial Neural Networks and Machine Learning–ICANN 2021: 30th International Conference on Artificial Neural Networks, Bratislava, Slovakia, September 14–17, 2021, Proceedings, Part IV 30* (pp. 372-383). Springer International Publishing. https://doi.org/10.1007/978-3-030-86380-7_30

Kotal, S. D., & Bhattacharya, S. K. (2020). Improvement of wind field forecasts for tropical cyclones over the North Indian Ocean. *Tropical Cyclone Research and Review*, *9*(1), 53-66. https://doi.org/10.1016/j.tcrr.2020.03.004

Chen, J., & Chavas, D. R. (2023). A Model for the Tropical Cyclone Wind Field Response to Idealized Landfall. *Journal of the Atmospheric Sciences*, *80*(4), 1163-1176. https://doi.org/10.1175/JAS-D-22-0156.1

Matyas, C. J. (2010). Associations between the size of hurricane rain fields at landfall and their surrounding environments. *Meteorology and atmospheric physics*, *106*, 135-148. https://doi.org/10.1007/s00703-009-0056-1

Yang, Y. H., Ying, M., & Chen, B. D. (2009). The climatic changes of landfall tropical cyclones in China over the past 58 years. *Acta Meteorol Sin*, *67*(5), 689-696.

Nie, X., Tan, H., Cai, R., & Gao X. (2023). Projection of the tropical cyclones landing in China in the future based on regional climate model. *Advances in Climate Change Research*, *19*(1), 23. https://doi.org/10.12006/j.issn.1673-1719.2022.064

Richman, M. B., & Leslie, L. M. (2012). Adaptive machine learning approaches to seasonal prediction of tropical cyclones. *Procedia Computer Science*, *12*, 276-281. https://doi.org/10.1016/j.procs.2012.09.069

Tan, J., Liu, H., Li, M., & Wang, J. (2018). A prediction scheme of tropical cyclone frequency based on lasso and random forest. *Theoretical and Applied Climatology*, *133*, 973-983. https://doi.org/10.1007/s00704-017-2233-3

Chen, Y. (2022). Research on Prediction of Typhoon Frequency in theNorthwest Pacific Based on Multi-model Integration (Master's thesis). Central South University. https://doi.org/10.27661/d.cnki.gzhnu.2022.001080

Chen, R., Zhang, W., & Wang, X. (2020). Machine learning in tropical cyclone forecast modeling: A review. *Atmosphere*, *11*(7), 676. https://doi.org/10.3390/atmos11070676

China Meteorological Administration Tropical Cyclone Data Center. 1952-2022. CMA Best Track Dataset. https://tcdata.typhoon.org.cn/zjljsjj.html

Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., Zhu, Y., & Chen, D. (2014). An overview of the China Meteorological Administration tropical cyclone database. *Journal of Atmospheric and Oceanic Technology*, *31*(2), 287-301. https://doi.org/10.1175/JTECH-D-12-00119.1

Lu, X., Yu, H., Ying, M., Zhao, B., Zhang, S., Lin, L., Bai, L., & Wan, R. (2021). Western North Pacific tropical cyclone database created by the China Meteorological Administration. *Advances in Atmospheric Sciences*, *38*, 690-699. https://doi.org/10.1007/s00376-020-0211-7

National Climate Center. 1951-2022. A Collection of 100 Climate System Indices. http://cmdp.ncc-cma.net/Monitoring/cn_index_130.php

Rong, X., Qin, W., & Wei, W. (2023). Typhoon Number Prediction Based on Three Machine Learning Algorithms. *Marine Forecasts*, *40*(5), 1-9.

Improved Satellite-Based Emergency Mapping through Automated Triggering of Processes

Martin Mühlbauer
German Aerospace Center
Martin.Muehlbauer@dlr.de

Monika Friedemann
German Aerospace Center
Monika.Friedemann@dlr.de

Johanna Roll
German Aerospace Center
Johanna.Roll@dlr.de

Torsten Riedlinger
German Aerospace Center
Torsten.Riedlinger@dlr.de

Fabian Henkel
German Aerospace Center
Fabian.Henkel@dlr.de

Lucas Angermann
German Aerospace Center
Lucas.Angermann@dlr.de

Mathias Böck
German Aerospace Center
Mathias.Boeck@dlr.de

Tobias Kaminski
German Aerospace Center
Tobias.Kaminski@dlr.de

Kayla Barginda
German Aerospace Center
Kayla.Barginda@dlr.de

ABSTRACT

For more than two decades, satellite-based emergency mapping (SEM) supports authorities and responders in the rapid assessment of disaster situations. Although automation has been advanced, particularly in the field of satellite image analysis, the standard SEM workflow of the Copernicus Emergency Mapping Service (CEMS) is mostly user-driven at crucial steps. The 2021 flood events in western Germany show, as an example, the strong influence of manual interaction on the timely provision of satellite-based crisis products to emergency managers. We examine where latencies occur in the CEMS workflow and show a concept for the automation of processes and the usage of primarily open web data for process triggering to overcome delays. Our assumption is that a combination of both components is key for a faster SEM workflow, which we discuss along further benefits and challenges. A prototypical information system demonstrates the practical applicability of the concept.

Keywords

Earth observation, disaster management, satellite-based emergency mapping, web data analysis.

INTRODUCTION

Large-scale natural disasters such as the severe floods in western Germany in 2021 (e.g. Ahr valley) or Pakistan in 2022, the extensive wildfires in Portugal in 2017, Spain in 2022, Chile in 2023 and 2024 or the devastating earthquake in the bordering region of Turkey and Syria in 2023 require an effective, fast and coordinated disaster management (DM). When a disaster strikes, rapid, timely situation assessment is key for an adequate operational planning, so that available resources can be used effectively, shown by the example of the Ahr valley flood (Holzheimer et al., 2022). Earth observation (EO) data products have proven highly valuable in providing an up-to-date overview of an affected region and valuable information on the extent and distribution of damage (Gstaiger et al., 2022). Satellite-based emergency mapping (SEM) services such as the Copernicus Emergency Management Service (CEMS) rapid mapping service (RMS) (CEMS, 2024) provide geospatial information on demand and fast in support of DM activities before, during or immediately following a disaster (Voigt et al., 2016).

In contrast to sudden-onset disasters like earthquakes, the availability of early warnings, e.g. for meteorological events, may enlarge the time window for an improved DM support with EO data. Recent undertakings to accelerate SEM in terms of delivery timeliness of the crisis information by utilizing early warning systems have proven effective (Wania et al., 2021), however the process remains user-driven at crucial steps. In this paper, we investigate possible enhancements for the SEM process and propose a concept for automation of manual steps to improve both, timely provision and accuracy of EO-based crisis information. Afterwards, we show how our approach is applied in a prototypical implementation as a proof-of-concept.

Our work is based on design science research (Hevner et al., 2004). The presented artifacts, an improved SEM process and an information system prototype, are the first results of an ongoing development process with practice and state-of-the-art research in mind. This is made possible in particular by the authors' position as employees of an RMS (Center for Satellite-Based Crisis Information (ZKI)) based at a research organization (German Aerospace Center (DLR)). Both the relevance and rigor of the artifacts can thus be iteratively ensured.

RELATED WORK

Since 2003, ZKI has established one of the first rapid mapping mechanisms that continued to evolve with concurrent advances in satellite technology to provide users with precise EO-based crisis information (Gähler, 2016). Among other RMS providers, DLR has contributed to the predecessor of the CEMS (Schneiderhan et al., 2010). Today, the standard SEM process of the CEMS starts with a user-driven activation, which on average is requested 20h after the user-defined event time. Upon activation, the state-of-the-art process follows the steps 1) satellite tasking (optional, see chapter “Satellite Acquisitions”) 2) image acquisition, 3) image delivery and 4) map product delivery including image analysis (Wania et al., 2021) (Figure 1).

The timeline for the severe Ahr valley flooding was reconstructed to illustrate this SEM process. Figure 2 shows early warnings by the European Flood Awareness System (EFAS, https://www.efas.eu) and the German National Meteorological Service (DWD, 2021a, 2021b) four, respectively, one day before the CEMS activation by the German Joint Information and Situation Centre (GMLZ). Upon the activation, cloud cover prevented the use of optical satellite imagery for damage assessment during the initial flooding days. The ZKI became active at request of the Bavarian Red Cross given the DWD warnings and by chance was able to provide timely aerial imagery (Holzheimer et al., 2022). An automatic flood detection based on Sentinel-1 radar satellite images was also published (ZKI, 2021). This case exemplifies the time delays in the current SEM workflow due to manual activation responsibilities and handover procedures. It is important to note that the disaster response cannot be reconstructed for a theoretical case in which the SEM would have been activated at an earlier stage.

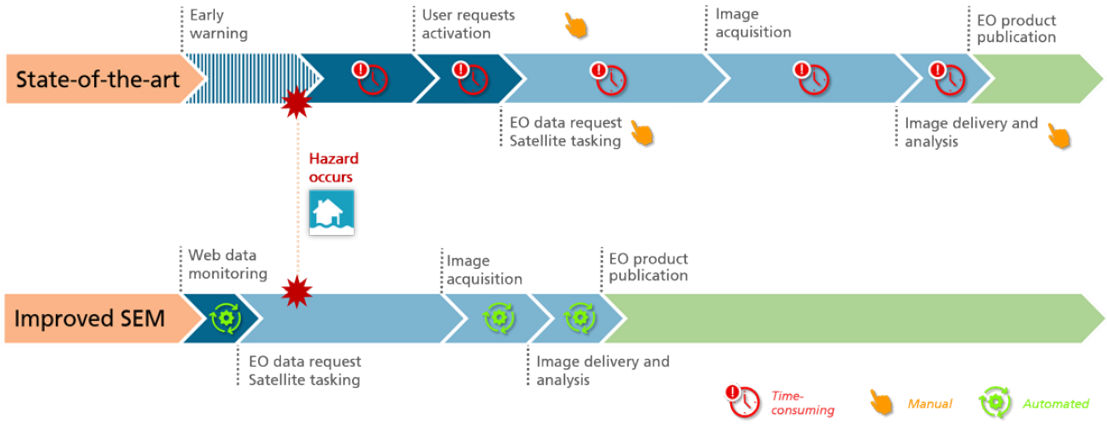


Figure 1. Comparison between state-of-the-art and proposed automated SEM workflow

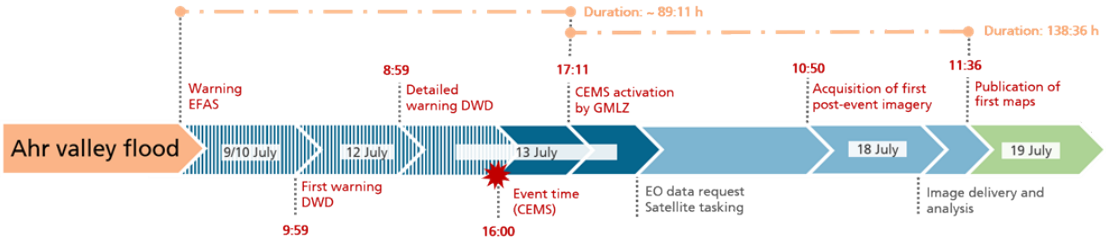


Figure 2. Condensed timeline of the RMS activation EMSR517 for Bad Neuenahr-Ahrweiler with time in UTC (based on CEMS, 2021a, 2021b; Weidinger, 2022)

Overall, we can identify several significant gaps in time delay 1) between the initial early warning and the SEM activation, being manually performed by an authorized user, and 2) between the activation and satellite tasking. Furthermore, 3) the latency between satellite sensing and image delivery (e.g. on Copernicus Open Access Hub, https://scihub.copernicus.eu) and 4) the time- and labor-intense manual and semi-automatic visual image interpretation delay the SEM process at later steps. Most attention has been given to optimizing the latter gap, which remains an area for ongoing improvements, such as fully automated flood detection (Wieland & Martinis, 2019; Bereczky et al., 2022), and burnt area derivation (Nolde et al., 2020; Knopp et al., 2020).

Considering the first two gaps, mental workload and initial uncertainties affect both users and SEM providers during disasters. Authorities must be aware of the event location to define the area of interest (AOI) for a SEM activation. End users are often unaware of possible and upcoming satellite data acquisitions, including Copernicus and very high-resolution (VHR) satellites, and the time until map product availability. SEM service providers are engaged in a time-consuming search and coordination for suitable satellite data sources. Reducing manual activation time delay is crucial given past delays in RMS activation of, for instance, 36h after disasters in Bosnia and Herzegovina (CEMS, 2014). Since 2016, EFAS has been linked to the CEMS SEM process to anticipate satellite tasking for riverine floods. Wania et al. (2021) found that EFAS pre-tasking reduced delivery time for the first map product from 28:47 h to 16:05 h on average considering 14 flood events in Europe from 2016 to 2020. Similarly, Ajmar et al. (2019) proposed to include early warning alerts like those of the Global Disaster Alert and Coordination System (GDACS, https://www.gdacs.org), which also integrates tsunami models to reduce acquisition times by up to one day (GDACS, n.d.).

THE PROPOSED CONCEPT

The SEM workflow described above reveals that manual interaction and latency are the main reasons for a delayed provision of EO-based data products for DM, particularly the triggering step and the search for suitable satellite data sources. There are already approaches in place that shorten the time for triggering, but these are limited to specific natural events, and in some cases are not commonly accessible. Thus, we propose an approach that substitutes manual interaction and latency with a) the automation of processes and b) the usage of primarily open data from the WWW. Our assumption is that combining both components is key for an accelerated and therefore substantially improved SEM workflow.

AOI Definition

Figure 1 shows our proposed workflow for automating the initial steps of a SEM activation, which requires two types of input data: a) data that allows us to determine an AOI and b) data that contains information on future satellite data acquisitions. In our approach, the AOI is defined as an automatically derived area in which an event is likely to happen in the near future or has just happened. An event can be anything that might require an observation with remote sensing technologies, such as a natural or man-made disaster. While the AOI is currently determined manually by the entity that requests the satellite data, we propose a process that automatically retrieves the AOI by monitoring crisis-related data from the WWW. Monitoring involves continuously analyzing heterogeneous data to detect spatiotemporal anomalies in a given context.

Sometimes web data contain such anomalies implicitly. In case of natural disasters, for example, numerous national and international entities issue official alert messages already including an AOI (e.g. MeteoAlarm, https://www.meteoalarm.org). For predictable events, such as storms, these alerts can be early warnings. For sudden-onset disasters, such as earthquakes, these alerts are usually published in a short time frame after the event.

Another example of publicly available spatiotemporal anomalies are event databases like the Global Database of Events, Language, and Tone (https://www.gdeltproject.org). These databases may contain conflict-related events, such as air raids. Technically, we can directly use the spatiotemporal information from alerts and event databases for building an AOI. Yet, trustworthiness of the sources must be considered. Warnings issued by widely accepted institutions such as GDACS can be regarded as very reliable. We can trust the expertise of the specialists and integrate the provided AOIs into our process. However, sometimes there are no official warnings or up-to-date event databases for the region or the event type we are interested in. As an example, quickly emerging refugee camps or natural disasters in countries without warning infrastructure can be mentioned. In such cases, or when the information derived from alerts and event databases is not satisfactory, other data sources should be considered, for example media monitors (e.g. Europe Media Monitor, https://emm.newsbrief.eu) and social media platforms. Here, the challenges are data access, data reliability and the fact that the spatiotemporal anomaly detection has to be performed by ourselves. An automated event detection approach (see Li et al. (2022) for examples) can be applied to turn unstructured textual data into well-structured events that describe the location, the time and the type of a disaster (Resch et al., 2018; Havas & Resch, 2021).

In summary, we can assume that in certain situations only one data source might be sufficient for the definition of an AOI. Even more, this applies to situations where a) the size of the footprint of the satellite is much bigger than the AOI, b) the satellite data are available free of charge and c) the satellite data will be analyzed automatically. In other situations, we cannot rely on only one data source. For instance, an official alert might cover a vast administrative entity. However, an RMS provider is primarily interested in the situation in populated areas. In such cases, we propose the fusion of as many data sources as necessary to determine a sensible AOI. In this scenario, together with the official alert(s), we could use a population distribution dataset and social media analysis for refinement (Lorini et al., 2019). The efforts pay off particularly when a) the size of the footprint of the satellite is smaller or approximately of the same size as the AOI, b) the satellite data are only available commercially at a high price and c) the satellite data will be analyzed manually.

Satellite Acquisitions

Besides the definition of AOIs for specific events, information on satellite acquisitions is essential in our approach. We have to distinguish three different satellite categories depending on the availability of satellite data. 1) Some EO satellites collect data permanently. The data are made available free of charge on platforms on the WWW. Besides the actual imagery, acquisition planning data are sometimes published on the WWW. Hence, we know when and where a satellite will acquire data in the coming days. Examples for this category are the Copernicus Sentinel (e.g. Sentinel-2, see European Space Agency, n.d.) or Landsat missions. 2) Some EO satellites offer a direct data broadcast. These satellites still collect data permanently. However, ground stations can directly receive the data when the satellite passes over and before the data is published on official platforms. Alongside the sometimes publicly available acquisition planning data, information from and about ground stations is required to assess data availability. Examples of this category are Sentinel-1, Aqua, Terra and Suomi NPP. 3) Some EO satellites do not collect data on a permanent basis. Instead, these satellites have to be tasked by an authorized entity. For acquisition planning, a ground station has to uplink information to the satellite, such as where and when data have to be acquired. Examples of this category are commercial VHR optical satellites (e.g. WorldView-3) and radar satellites (e.g. TerraSAR-X).

Data Intersection

For satellites from category one, as defined in chapter “Satellite Acquisitions”, each of the derived AOIs can be intersected with the potential sensing areas. No manual interaction is required and all the necessary data can be downloaded from the WWW. For satellites from category two, we additionally have to consider the direct acquisition capacities of ground stations. For satellites from category three an interface with commercial satellite operators has to be established in order to ensure that the requested AOIs can be considered in acquisition planning. The overall goal remains the same in all three cases: We want to know which satellite data will be available for our AOIs at what time. This information is of great value to decision makers in crisis situations based on experience gathered from previous ZKI activations and research projects, e.g. EU FP7 PHAROS (Community Research and Development Information Service (CORDIS), 2023) and EU H2020 HEIMDALL (CORDIS, 2022). At this point of the workflow, users can quickly be informed by means of decision proposals, for instance short email notifications, about their options for their customized satellite-based product. Thus, decision proposals are a further means to strengthen situational awareness (Endsley, 1995). While our system is intended to function fully automatically until this step, the relevance of decision proposals becomes evident if the acquisition of possibly expensive commercial data is proposed.

When EO analyses are based on publicly available data, further automation is possible and intended. In our approach, as soon as the EO data for the derived AOI are available, automatic thematic processors would instantly start analyzing these data. After a manual expert-based quality control, the EO product can be delivered to users. Besides the important EO product validation, we envisage an equally important step of AOI validation. Above we have shown that AOI derivation might be based on our own spatiotemporal anomaly detections. In such cases, we have to ensure that thresholds are calibrated correctly in order to constantly improve the quality of derived AOIs.

PROTOTYPE AND EXAMPLE

In order to demonstrate the feasibility of our concept we have implemented an experimental application based on disaster alerts from various official international (e.g. GDACS) and national (e.g. German DWD) sources as well as acquisition data of Sentinel-1A, Sentinel-2A, Sentinel-2B und Landsat-8. The prototype automatically collects the alerts from the WWW and calculates when the potential sensing area of a satellite intersects with an AOI. The results are saved in a database that can be queried by a graphical user interface for visualization purposes. Additionally, a dissemination service automatically generates textual notifications with current disaster events and matching satellite acquisition options.

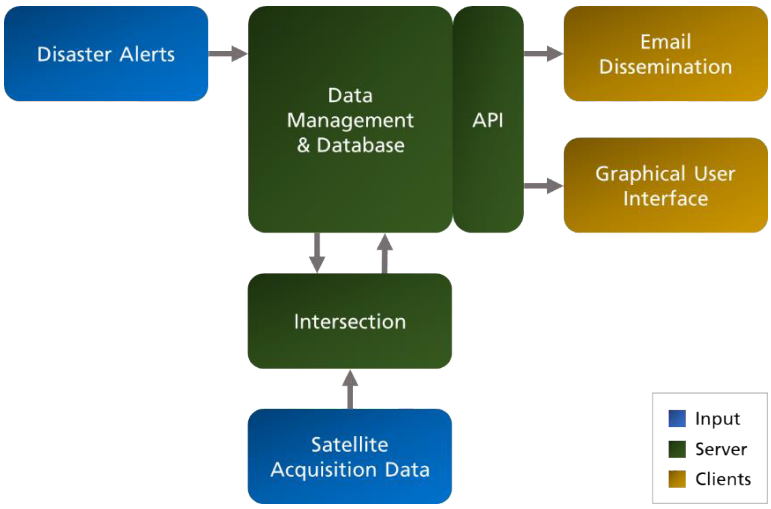


Figure 3. Major components of the prototype. The arrows indicate the direction of the data flow.

Figure 3 illustrates the major components of the developed prototype and their interactions. The prototype is designed as a client-server web application (Gallaugher & Ramanathan, 1996). The server comprises a data management component, an intersection component, a database and an API. The data management component is built with Python. Its tasks are, for instance, the download of alerts and the triggering of the intersection component. The spatiotemporal intersection component is a Java-based extension of the operational PSM (Processing System Management) (Böttcher, 2001) deployed at DLR's ground station for satellite data reception in Neustrelitz, Germany. It selects those satellite acquisitions within the next 7 days that spatially and temporally overlap with the AOIs from the alerts and returns the results to the data management. The communication between data management and intersection runs via gRPC (gRPC Remote Procedure Calls) (https://grpc.io). Next, the results are stored in a PostgreSQL database. A GraphQL API (https://graphql.org) built with Hasura (https://hasura.io) delivers the requested results to clients. The server part runs automatically at regular intervals, e.g. once per hour. A dissemination service written in Python generates brief email notifications based on the intersection results. For a more detailed insight into the collected and generated data, there's a web-based graphical user interface with an interactive map and various sorting and filter options. The graphical user interface is implemented with DLR's open source software for geoscientific web applications UKIS (Environmental and Crisis Information Systems) (Muehlbauer, n.d.; https://github.com/dlr-eoc/ukis-frontend-libraries).

The following screenshots of the prototype's graphical user interface show exemplarily how our system worked in the context of a predicted flood in northern Germany in November 2023. The "Warnings" view in Figure 4 gives an overview on all disaster alerts currently registered in the system. After selecting one particular alert, the system displays the AOI together with the matching satellite acquisitions. See "Detail" view in Figure 5.

Due to automation, the prototype was able to generate these results within minutes after the official warning was issued. As the underlying alert was an early warning, we could predict which satellite data would be available at what point of time before the actual event had happened.

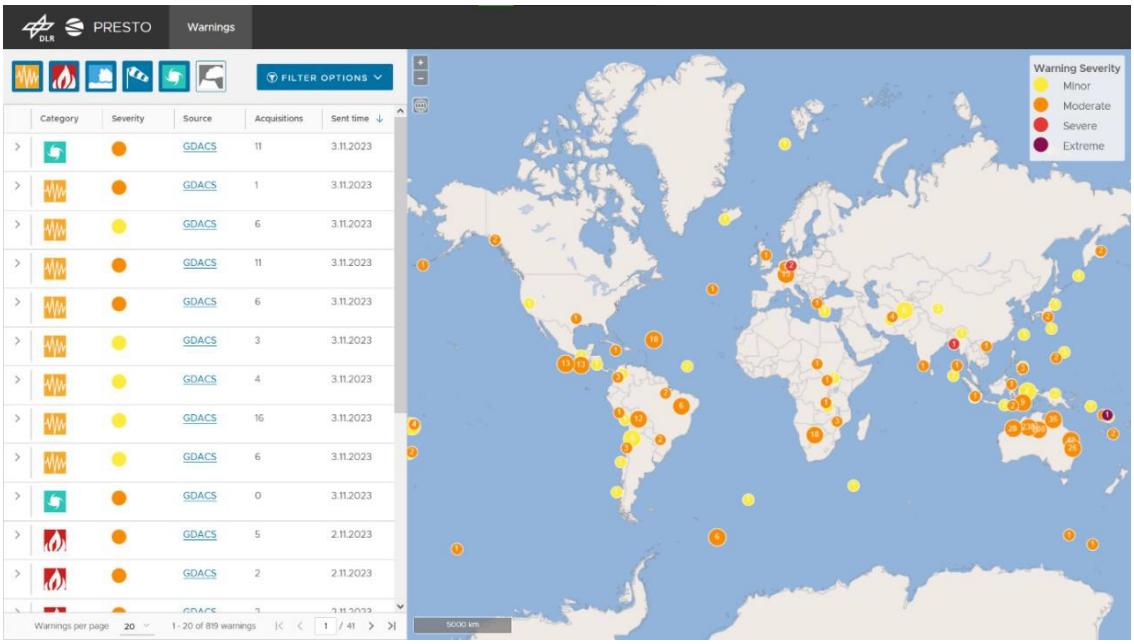


Figure 4. "Warnings" view of the prototype: The list at the left shows alerts for natural disasters with some of their attributes. The column "Acquisitions" contains the number of possible satellite data acquisitions for the respective AOI. The list can be filtered (here: source equals "GDACS") and ordered (here: time stamp descending) by the user. The map at the right shows the locations of the alerts, with colors indicating the severity of an alert.

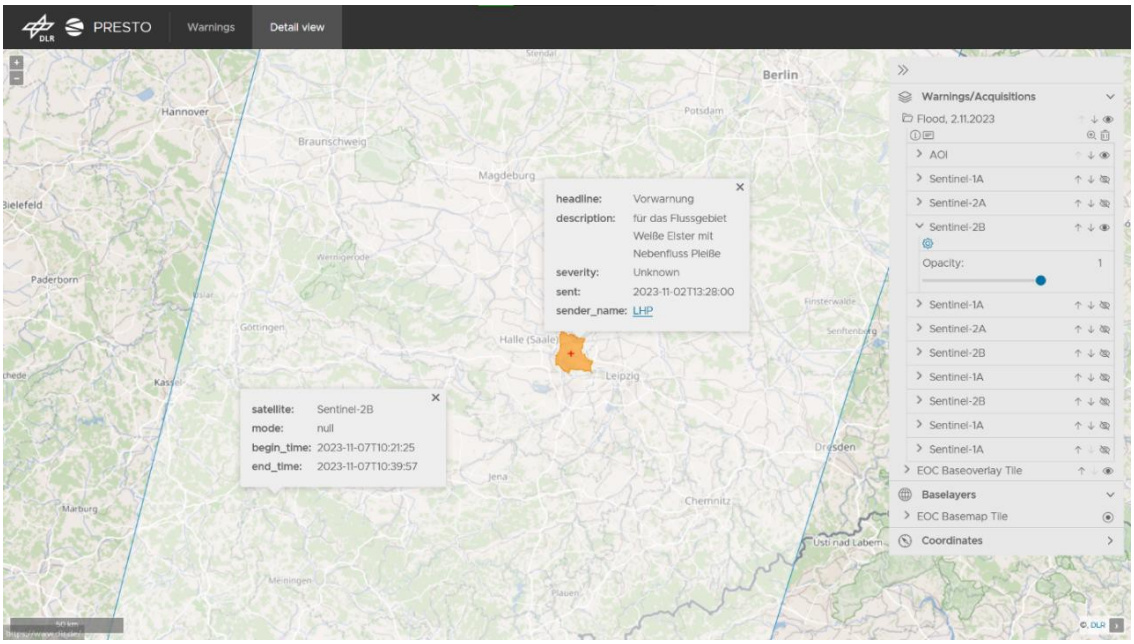


Figure 5. "Details" view of the prototype: The layer control element at the right contains entries for the AOI as well as all matching satellite acquisitions. A click on the AOI displays information such as the source of the alert and the original alert description. A click on an acquisition segment shows attributes such as the satellite name or the sensing time(s).

DISCUSSION

We consider our approach as a possible advancement of well-established SEM workflows addressing different kinds of natural or man-made disasters. Proven SEM process steps are combined with automation and primarily open data in order to satisfy the prevalent requirement of users of SEM products: faster information and product availability. The main objective is to improve the timeliness, but we also anticipate improved quality in the final EO products. If data interpretation cannot be fully automated, an EO expert will have additional time for analysis. Also, alerts, media articles etc. used to trigger the workflow can enrich the final product. Our approach improves the user's situational awareness by automatically generating decision proposals regarding EO data and product availability. This can also comprise crucial information such as the expected cloud coverage. Besides the user's benefits, there are advantages on the RMS provider's side. The monitoring of trigger data on a continental or global scale can lead to useful products that would not have been possible with human triggers alone. Cost is another factor. As long as no commercial data acquisition is involved, the workflow is based on free input data. Of course, we acknowledge the significant challenges associated with automation and open data. We heavily rely on trusted data sources such as official alerts. Yet, even official alerts might contain errors. Other possible data sources for the AOI generation, particularly media and social media, must be treated even more carefully. We try to eliminate as many inaccuracies as possible with a fusion of a reasonable amount of data sources in the respective use case. To achieve this fusion, the heterogeneous data sources first could be aggregated geometrically e.g. by incorporating the discrete global grid system H3 (https://h3geo.org). Afterwards, rule-based criteria can be applied to further validate and refine the AOI as much as possible (Roll, 2024). Moreover, a manual validation step for the AOI generation should be in place. The fusion approach also tries to absorb the loss of web data sources in course of time. Given that, for example, official alerts, events from online news and events from social media were used to define AOIs, the system would continue to function if social media were removed. Certainly, we cannot eliminate an inherent problem of satellite-based EO: the satellite most likely is not at the position over a disaster event or has just passed by. Latency is the result. Also, cloud coverage might further delay the assessment of optical imagery.

An automatic trigger can generate the decisive advantage in time, for example, by allowing data reception via direct broadcast. While automatically intersecting publicly available AOIs and acquisition plans is straightforward, including ground stations and commercial data providers may appear technically and administratively obstructive. However, an interface with DLR's ground station for satellite data reception in Neustrelitz, Germany could be established for our prototype. Finally, we want to emphasize the importance of the human factor in our approach. Crucial decisions will still – and have to – be made by humans based on their experience and transparent decision proposals. Expert knowledge from specialists at the disaster site is, once available, of utmost significance and we treat it as yet another – but very trustworthy – data source of the system.

SUMMARY AND FUTURE WORK

In this paper, we presented our vision of an enhanced SEM workflow as well as a first proof-of-concept implementation. We believe that our workflow can help to narrow down some gaps in the current state-of-the-art SEM. An accelerated process, better transparency, and more high-quality EO products can be achieved by automated processes and the combination of existing data. First experiments based on publicly available alerts and satellite acquisition data from DLR's ground station have demonstrated promising results: The prototype shows that satellite data availabilities can automatically be determined for areas potentially affected by a crisis. Our future work will focus on the transformation of our vision into a stable software that reliably supports future SEM activities. This includes a logic that considers the suitability of each satellite for each disaster type. Also, we aim at automatically generating more refined AOIs by aggregating available warning sources and filtering relevant attributes. Eventually, we plan to link our system to processing chains for satellite data in order to automate a further process step. Not only the 2021 Ahr valley flood in Germany has shown that an improvement of existing processes is relevant and possible.

ACKNOWLEDGMENTS

We certify that there is no actual or potential conflict of interest in relation to this article. This work is part of the two projects TEMA (Trusted Extremely Precise Mapping and Prediction for Emergency Management) and IFAS (Indicator Monitoring for Early Acquisition of Innovative Satellite Sensors in Natural Disasters). TEMA receives funding from the European Union's Horizon Europe research and innovation program under grant agreement No 101093003. IFAS is funded by the Bavarian Ministry of Economic Affairs, Regional Development and Energy in the Bavarian Collaborative Research Program (BayVFP) framework.

REFERENCES

Ajmar, A., Annunziato, A., Boccardo, P., Tonolo, F.G., & Wania, A. (2019). Tsunami Modeling and Satellite-Based Emergency Mapping: Workflow Integration Opportunities. *Geosciences*, 9(7), 314. https://doi.org/10.3390/geosciences9070314

Bereczky, M., Wieland, M., Krullikowski, C., Martinis, S. & Plank, S. (2022). Sentinel-1-Based Water and Flood Mapping: Benchmarking Convolutional Neural Networks Against an Operational Rule-Based Processing Chain. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 2023-2036. https://doi.org/10.1109/JSTARS.2022.3152127

Böttcher, M., Reißig, R., Mikusch, E. & Reck, C. (2001). Processing Management Tools for Earth Observation Products at DLR-DFD. *Procedings of DASIA*, France.

Community Research and Development Information Service. (2022). *Project on a multi-hazard open platform for satellite based downstream services*. European Union, CORDIS EU research results. https://cordis.europa.eu/project/id/606982

Community Research and Development Information Service. (2023). *Heimdall – Multi-hazard cooperative management tool for data exchange, response planning and scenario building*. European Union, CORDIS EU research results. https://cordis.europa.eu/project/id/740689

Copernicus Emergency Management Service. (2024). *Copernicus Emergency Management Service – Mapping*. https://emergency.copernicus.eu/mapping

Copernicus Emergency Management Service. (2021a). *EMSR517: Flood in Western Germany*. European Union. https://emergency.copernicus.eu/mapping/list-of-components/EMSR517

Copernicus Emergency Management Service. (2021b, 16 July). *The Copernicus Emergency Management Service forecasts, notifies, and monitors devastating floods in Germany, Netherlands, Belgium and Switzerland*. https://emergency.copernicus.eu/mapping/ems/copernicus-emergency-management-service-forecasts-notifies-and-monitors-devastating-floods

Copernicus Emergency Management Service. (2014). *EMSR087: Floods in Bosnia and Herzegovina*. European Union. https://emergency.copernicus.eu/mapping/list-of-components/EMSR087

Deutscher Wetterdienst [@DWD_presse]. (2021a, 13 July). *#Update zur #Starkregen-/#Dauerregen-Lage: Im Umfeld der Eifel ist mit den höchsten Niederschlagsmengen zu rechnen, daher wurden die #Warnungen in* [Post]. X. https://twitter.com/DWD_presse/status/1414872018005241857

Deutscher Wetterdienst [@DWD_presse]. (2021b, 12 July). *Besonders Teilen des Westens drohen in den nächsten Tagen große Regenmengen. Im Südwesten muss man sich mancherorts bereits in der* [Image attached] [Post]. X. https://twitter.com/DWD_presse/status/1414524803273400323

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, 37(1), 32-64. http://dx.doi.org/10.1518/001872095779049543

European Space Agency. (n.d.). *Acquisition Plans.* https://sentinel.esa.int/web/sentinel/copernicus/sentinel-2/acquisition-plans

Gähler, M. (2016). Remote sensing for natural or man-made disasters and environmental changes. *Environmental applications of remote sensing*, 309-338.

Gallaugher, J. M. & Ramanathan, S. C. (1996). Choosing a Client/Server Architecture. *Information Systems Management*, 13(2), 7-13. https://doi.org/10.1080/10580539608906981

Global Disaster Alert and Coordination System. (n.d.). *Models – Earthquake*. Retrieved February 12, 2024, from https://www.gdacs.org/Knowledge/models_eq.aspx.

Gstaiger, V., Merkle, N., Rosenbaum, D., Wieland, M., Lechner, K., & Kippnich, U. (2022). Aus dem All und aus der Luft frisch auf den (Lage-)Tisch: Der Nutzen von Luft- und Satellitendaten für die Lageerfassung. *Im Einsatz*, 56-61.

Havas, C., & Resch, B. (2021). Portability of semantic and spatial–temporal machine learning methods to analyse social media for near-real-time disaster monitoring. *Natural Hazards*, *108*(3), 2939-2969. https://doi.org/10.1007/s11069-021-04808-4

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 75-105.

Holzheimer, E., Kippnich, U., Kippnich, M., Lechner, K., & Wieland, M. (2022). Erkundung im Ahrtal mit Unterstützung von Verfahren der Künstlichen Intelligenz. *Die Flut im Juli 2021. Erfahrungen und Perspektiven aus dem Rettungsingenieurwesen und Katastrophenrisikomanagement*, 1, 22-26.

Knopp, L., Wieland, M., Rättich, M., & Martinis, S. (2020). A Deep Learning Approach for Burned Area Segmentation with Sentinel-2 Data. *Remote Sensing*, 12(15), 2422. https://doi.org/10.3390/rs12152422

Li, Q., Chao, Y., Li, D., Lu, Y., & Zhang, C. (2022). Event Detection from Social Media Stream: Methods, Datasets and Opportunities. *2022 IEEE International Conference on Big Data (Big Data)*, Osaka, Japan, 3509-3516. https://doi.org/10.1109/BigData55660.2022.10020411

Lorini, V., Castillo, C., Peterson, S., Rufolo, P., Purohit, H., Pajarito, D., De Albuquerque, J.P., & Buntain, C. (2021). Social media for emergency management: Opportunities and challenges at the intersection of research and practice. *Proceedings of the 18th International Conference on Information Systems for Crisis Response and Management*, USA, 772-777.

Muehlbauer, M. (n.d.). *UKIS – Environmental and Crisis Information Systems*. German Aerospace Center, Earth Observation Center. Retrieved May 14, 2024, from https://www.dlr.de/de/eoc/ueber-uns/deutsches_fernerkundungsdatenzentrum/georisiken-und-zivile-sicherheit/informationssysteme-und-geomatik/ukis-umwelt-und-kriseninformationssysteme

Nolde, M., Plank, S., & Riedlinger, T. (2020). An Adaptive and Extensible System for Satellite-Based, Large Scale Burnt Area Monitoring in Near-Real Time. *Remote Sensing*, 12(13), 2162. https://doi.org/10.3390/rs12132162

Resch, B., Usländer, F., & Havas, C. (2018). Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. *Cartography and geographic information science*, *45*(4), 362-376. https://doi.org/10.1080/15230406.2017.1356242

Roll, J., Schneibel, A., Mühlbauer, M. & Gähler, M. (2024, March 7-8). *Aggregation of heterogeneous data sources for the early acquisition of satellite images in case of flood events* [Book of abstracts]. Discussing Geographic Method(ologie)s: Integration, Experimentation, and Innovation, Innsbruck, Austria. https://fileshare.uibk.ac.at/f/ebef338271b44ca59816

Schneiderhan, T., Gähler, M., Kranz, O., & Voigt, S. (2010). Insights to the Emergency Mapping Service within the GMES project SAFER-Highlights, main achievements and challenges. *Conference Proceedings of the Living Planet Symposium*, Bergen, Norway, 1-5.

Voigt, S., Giulio-Tonolo, F., Lyons, J., Kučera, J., Jones, B., Schneiderhan, T., Platzeck, G., Kaku, KK., Hazarika, M.K., Czaran, L., Li, S., Pedersen, W., James, G.K., Proy, C., Muthike, D.M., Bequignon, J., & Guha-Sapir, D. (2016). Global trends in satellite-based emergency mapping. *Science*, 353(6296), 247-252. https://doi.org/10.1126/science.aad8728

Wania, A., Joubert-Boitat, I., Dottori, F., Kalas, M., & Salamon, P. (2021). Increasing Timeliness of Satellite-Based Flood Mapping Using Early Warning Systems in the Copernicus Emergency Management Service. *Remote Sensing*, 13(11), 2114. https://doi.org/10.3390/rs13112114

Weidinger, A. (2023, June 6). *Was ist in der Flutnacht passiert? – Ein Protokoll*. Südwestrundfunk SWR. https://www.swr.de/swraktuell/rheinland-pfalz/flut-rekonstruktion-ahrtal-protokoll-100.html

Wieland, M. & Martinis, S. (2019). A Modular Processing Chain for Automated Flood Monitoring from Multi-Spectral Satellite Data. *Remote Sensing*, 11(19), 2330. https://doi.org/10.3390/rs11192330

Zentrum für satellitengestützte Kriseninformation. (2021, July 15). *Unwetter in Nordrhein-Westfalen und Rheinland-Pfalz, Deutschland*. German Aerospace Center (DLR), Earth Observation Center. Retrieved February 14, 2024, from https://activations.zki.dlr.de/en/activations/items/ACT152.html

SILVANUS Dashboard for Wildfire Management

Dominika Grunwald
ITTI Sp. z o.o., Poznań, Poland
dominika.grunwald@itti.com.pl

Marcin Wawrzyniak
ITTI Sp. z o.o., Poznań, Poland
mwawrzyniak@itti.com.pl

Marcin Przybyszewski
ITTI Sp. z o.o., Poznań, Poland
mprzybysz@itti.com.pl

Emil Gatial
Institute of Informatics, Slovak Academy of Sciences, Slovakia
emil.gatial@savba.sk

Zoltan Balogh
Institute of Informatics, Slovak Academy of Sciences, Slovakia
balogh@savba.sk

Sepideh Hassankhani-Dolatabadi
Institute of Informatics, Slovak Academy of Sciences, Slovakia
sepideh.ui@savba.sk

ABSTRACT

The world is constantly exposed to the negative effects of wildfires, which occur for various reasons, often beyond human control. In such a case, it is highly advisable to develop appropriate countermeasures that will comprehensively approach the problem of wildfires. In response to these needs, the SILVANUS project was created. The goal of SILVANUS is to use new and existing technologies in the field of fire management, such as mobile applications for citizens, dedicated user interfaces or special applica tions for training firefighters in VR. The set of these tools is intended to support Integrated Fire Management. In this paper, we propose a web -based interface which facilitates crisis management during fires, especially wildfires. The Dashboard is an interactive map for monitored areas, with layers corresponding to different sources of data. It also supports communication between citizens and firefighters through mobile applications and social media. The main objective of the system is to help with decision making process so that the user will ultimately decide whether data is relevant or not. The solution is directed mainly at firefighters, analysts, and public administration entities. The presented Dashboard is a work in progress.

Keywords

Dashboard, wildfire, decision support system, map.

INTRODUCTION

Climate changes, voluntary or involuntary human activities, natural causes like lightning, are main reasons of wildfires that have a negative ecological and socio-economic impact on different countries. “Every year about 45000 forest fires occur in southern Europe, burning approximately 0.5 million hectares of forests and other rural lands.” (Moreira et al., 2011). Forests in Europe cover over 35% of total land area (FOREST EUROPE, 2020) and are responsible for many ecosystem services necessary for human we ll-being such as purifying the air, filtering water supplies, controlling floods and erosion, and many more. Unfortunately, when facing wildfires, forests are losing their ability to adapt and become more vulnerable to other dangers (Forzieri et al., 2021). We distinguish

controlled and unplanned forest fires. The purpose of a controlled burning, among other things, is to reduce the risk of uncontrolled fire and to initiate it intentionally to support the forestry and agricultural sectors or to reduce greenhouse gas emissions (Grebner et al., 2013). An unplanned forest fire is unwanted wildland fire which affects an increasing number of countries, therefore the need to take remedial actions is also growing.

While researching countermeasures to prevent the increasing number of wildfires, more and more authorities are turning to IT solutions. They are flexible, relatively easy to use and low cost, and can potentially be customized with information the user considers most necessary (Hunter & Westerman, 2009). The prevention and management of forest fires, restoration of areas destroyed by fire are the topics that the Horizon 2020 -funded project SILVANUS addresses. It aims to create a stable forest management platform that is resistant to climate change. It is intended not only to effectively prevent and suppress uncontrolled and dangerous forest fires, but also to help plan prescribed burning. As part of the project, SILVANUS uses the expertise of specialists in t he areas of environmental, technical and social sciences. Scientists and engineers are supporting civil protection authorities by helping to effectively monitor forest resources, assess biodiversity, create accurate fire risk indicators and promote safety regulations among local fire-affected communities through awareness campaigns. The project lasts 42 months and, with the involvement of 49 partners from around the world, is to develop the best possible solution in the fight against wildfires (SILVANUS, 2021).

One of the tools developed as part of the SILVANUS project is the Dashboard. It is an interactive panel that represents various data graphically, often through the use of charts, tables, maps, and other visual elements. The main purpose of this interface is to facilitate timely and easy user access to important information, as well as to enable monitoring of key indicators or metrics (Few, 2013). The SILVANUS Dashboard consists mainly of a map and corresponding layers intended to assist relevant authorities in making important decisions. The platform obtains information from various sources, such as robots, drones, computer analysis and even from residents who have the Citizen Engagement Application installed on their phones, which is another tool of the SIL VANUS project prepared to increase citizens' awareness about forest fires. The aim of this article is to provide information about the Dashboard prepared for the SILVANUS Project.

The article is structured as follows: the first two sections outline the context of the problem and present the current state of knowledge in the field of fire crisis management. The next section defines the SILVANUS Dashboard system requirements, presents the system description and its architecture. Finally, conclusions are presented, summarizing the contribution of this work to solving the original problem and presenting plans for further works.

STATE OF THE ART

Dashboard is a useful tool which is being used and analyzed for emergency management in various areas to help with the decision support process. Mordecai & Kantsepolsky (2018) prepared a framework to improve management in extreme weather conditions. They propose a system to effectively use data for decision making processes through open to change deployment and visualization of data streams and indicators. The functiona lity was developed on the basis of clear system requirements, received from real end -user needs. Another example is the ELD-BS tool suite which, among other things, is a web-based dashboard developed to help with the administration work in the event of a disaster. It also supports the exchange of information between the authorized units (Hellmund et al., 2022).

Numerous works are also being undertaken to analyze the problem of forest fires and their effects. The Sanborn Map Company proposed a commercial solution called Forestry in which they focus mainly on mapping, land modeling and spatial analytics. This tool supports urban tree canopy management, heat island mitigation and the work of wildfire fighting agencies (The Sanborn Map Company, Inc., n.d.). A different platform developed by a company Technosylva is a software called Wildfire Analyst™ for real-time analysis and simulation of fire behavior, with particular emphasis on performance. The user receives various useful information in the form of charts, reports or GIS maps. It is also cloud based which makes it very flexible. Examples of three Technosylva dashboard views can be found in Figure 1. (Technosylva, Inc., n.d.).

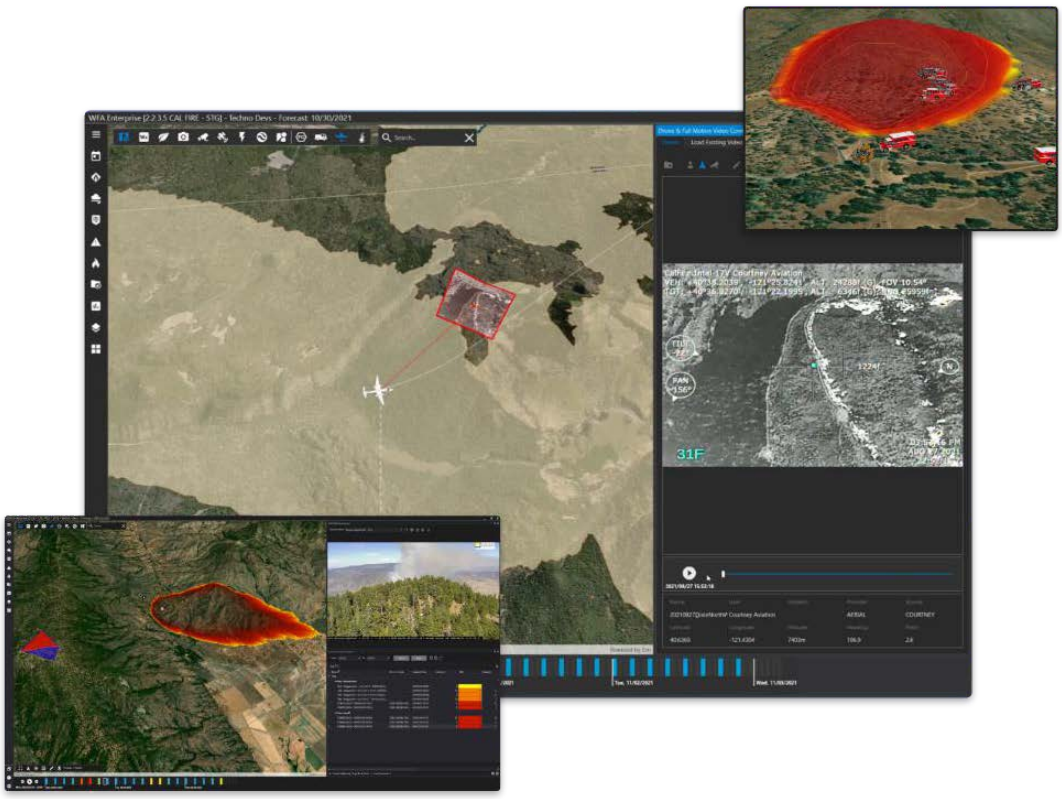


Figure 1. Technosylva dashboard (Technosylva, Inc., n.d.)

Another example is EMXSYS owned by Bruce Schubert and their CPS Wildfire Management Tool. It is free to use software for fire forecasting. Wildfire Management Tool uses the logical structure contained in Campbe **l** Prediction System, which is a fireground-specific language system that wildland firefighters learn from (Schubert, n.d.). An example of the EMXSYS dashboard view is shown in Figure 2.

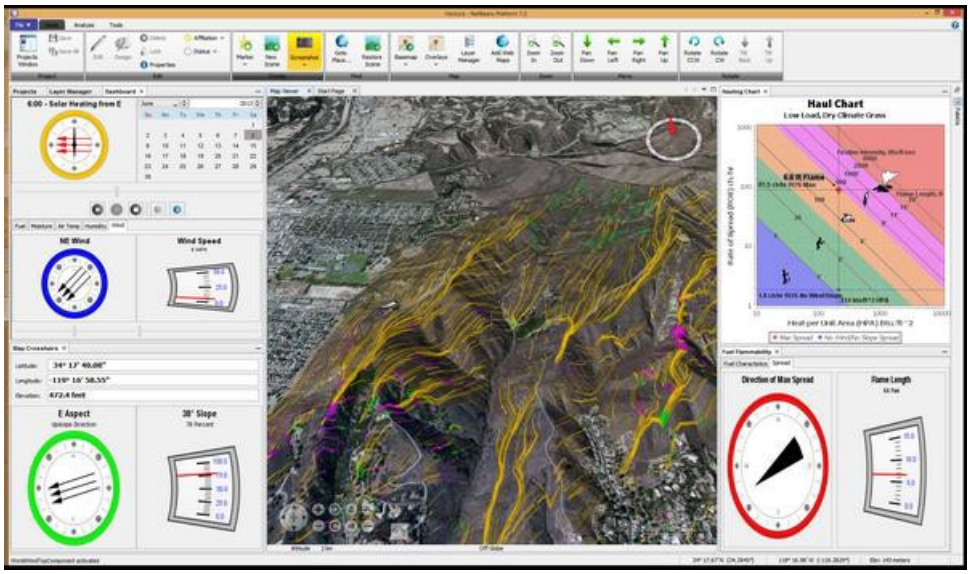


Figure 2. EMXSYS dashboard (Schubert, n.d.)

A comprehensive review of the literature shows that works on forest fires have been a topic often undertaken by specialists. One of the examples is the online GIS platform named AEGIS, which is a tool supporting the decision - making process in managing fire risk in Greece. It uses both spatial and non-spatial data to present the most useful information for the user about fires. The user can also conduct a fire simulation to understand how the fire will behave. Authors also used artificial neural networks to assess the risk of fire outbreaks, taking into account various parameters, learning methods, activation functions, data pre-processing techniques and network structures

(Battiston et al., 2019).

Another Horizon 2020 project called HEIMDALL also works on the problem of wildfires. It aims to enhance both immediate and long-term collaborative strategic planning at the regional level for the multiple stakeholders involved in disaster risk management a nd response. Their main goal was to create for decision-makers and other interested parties a platform that offers various tools to support crisis management. Many specialists from various specializations were engaged in the project. HEIMDALL platform predicts the spread and behavior of fire under a variety of fire conditions (Battiston et al., 2019).

Despite the large number of already existing solutions in the field of fire crisis management, the need to deal with the problem of wildfires is still urgent. According to our research and experience - it is important to use existing data to create new tools. The solutions do not have to be competitive, but rather complementary. For this reason, work on the SILVANUS project was started. This platform aims to minimize the number of fires, detect them faster and to help manage forest restoration better.

SILVANUS DASHBOARD

SILVANUS Dashboard is a web-based interface which is created to provide useful information, enabling users to quickly understand the situation on the field to make informed and good decisions. The interface is designed to be used mainly by trained personnel e.g. firefighters, analysts or public administration representatives to facilitate Integrated Fire Management. It combines prevention and firefighting strategies, taking into account social, economic, cultural and ecological analyses. Its goal is to minimize losses caused by fires and maximize the benefits that may come from them. In addition, it also covers four stages of crisis management: mitigation, preparedness, response and recovery. To use the Dashboard, a person must successfully undergo training conducted by a specialist. They also must demonstrate an understanding of fires, forest fires, legal regulations and how firefighting units operate in the served area.

To understand the needs of future users and to decide on the technology, contributors asked for opinion partners, firefighters and specialists through meetings, dashboard demonstrations and participation in pilots organized as part of the project. The whole process is shown in Figure 3.

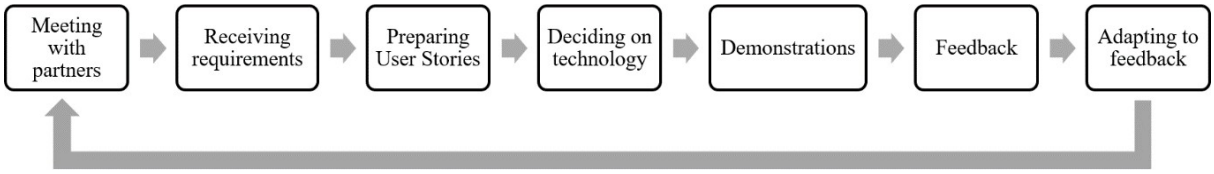


Figure 3. Process of developing Dashboard functionalities

The requirements collected at all meetings and levels were divided into functional and non -functional. They describe the constraints with which the system should perform its functions. All requirements are presented in Table 1 together with their descriptions.

Table 1. SILVANUS Dashboard functional and non-functional requirements

SRS-ID	Requirement	Description
SRS-F-1	Log in/out	A user must be able to log in/out to the existing account.
SRS-F-2	Displaying Layers	A user should be able to turn on/off the GIS layer while being in a map view.
SRS-F-3	Displaying Notifications	When the system gets an update about a new event the user must be notified about it on the map and on the notification bar.
SRS-F-4	Displaying right information drawer	When the user sees the new event on the map, he will be able to interact with it to see more information about it.
SRS-F-5	Displaying left information drawer	The left drawer must display layers, pin list, map typer, filter options.
SRS-F-6	Filtering	A user must be able to filter different layers by date and time range and different variables (depending on the layer).
SRS-F-7	Zooming in/out	A user must be able to zoom in/out the map.

SRS-F-8	Displaying Smoke/fire detection layer	A user must be able to browse pins with smoke/fire detection on the map.
SRS-F-9	Displaying Social Media Fire Events layer	A user must be able to browse pins from social media posts such as platform X with information about probable fires.
SRS-F-10	Displaying ID Social Media layer	The Dashboard must display a layer with area scope (set of coordinates) with information about it from social media in 3+ different languages.
SRS-F-11	Displaying Unmanned Ground Vehicles layer (UGV)	The Dashboard must display a set of pins on a map to show the route of UGV with photos and information gathered.
SRS-F-12	Displaying ID DSS layer	The Dashboard must display a layer with area scope (set of coordinates) with analytical information about it in different time ranges.
SRS-F-13	Displaying Fire Spread Simulation layer	The Dashboard must display a fire spread simulation as a polygon (set of coordinates).
SRS-F-14	Sending warnings	The user must be able to send warnings to citizen using Citizen Engagement Mobile App after clicking on a map and filling out the form.
SRS-F-15	Mapping a fire	The user must be able to draw a polygon on a map which will mirror the ongoing fire.
SRS-NF-1	Security	The Dashboard must be secured against unauthorized access.
SRS-NF-2	Performance	The Dashboard must be able to handle the required number of users with a constant flow of new information from the backend without compromising performance.
SRS-NF-3	Maintenance	The Dashboard should be easy to maintain and update.
SRS-NF-4	Simplicity	The interface would be used by people who have different knowledge in IT solutions so it should be easy to use.
SRS-NF-5	Verification	The Dashboard must meet with all applicable laws, terms and regulations, including Horizon 2020 terms (for example privacy by design).
SRS-NF-6	Reliability	The interface must be reliable and compliant with user requirements.
SRS-NF-7	Availability	The Dashboard must be available during system maintenance.

Main goals for this platform are to be transparent, intuitive and tailored to the needs of future users. After analyzing the collected requirements, a list of basic dashboard functionalities was created which are presented in Table 2.

Table 2. SILVANUS Dashboard functions

Function	Description
Layers	All data gathered from partners are displayed as layers in the form of highlighted areas, a collection of pins or a slider to show different images.
Pins	A pin is a feature to show a specific location. The user can interact with the pin by clicking on it to see more information about the event.
Area scope	This is a layer where the information concerns an area, i.e. a set of coordinates. The user can interact with the area by clicking on it to see more information.
Zoom in/out	The Dashboard allows to zoom in/out the map for better

	understanding by using a scroll on the mouse.
Notifications	When new data is delivered to the Dashboard, the user receives a notification. The user can delete the notification.
Filtering	All pins layers can be filtered by date and time.
Log in/out	Users can log in/out using an authorized account prepared by the admin.
Slider	The slider is used to quickly change the displayed graphics on the map (heatmap, fire spread simulation).
Drawing on the map	The user can specify an area on the map by placing points on the map. It can be edited or deleted.
Sending a poll to citizens	The Dashboard allows communication with Citizens Engagement App. The user of a Dashboard can send a warning or an announcement for specific area and time.
Map types	The map can be displayed in different types (satellite, street map).

Technical Information

This section describes the technologies used in the Dashboard interface and the general reasoning behind usage of said solutions. There is also a description of the data flow in the Dashboard as well as general architecture of integration of Dashboard with rest of SILVANUS through SAL (Storage Abstraction Layer).

Technologies

The core of the interface is written in ReactJS, which is one of the more recognizable, opensource libraries for building user interfaces based on components - standardized, interchangeable building blocks of UIs. It is being constantly developed and supported by a wide community of individual developers and companies. The library being open source ensures there will not be additional costs hidden in solution for SILVANUS being used and further developed in the future. Moreover, a wide variety of places where ReactJS is being currently used (ie. Facebook, Instagram, Netflix) can guarantee that it will be constantly improved and updated in the foreseeable future. These factors were considered as well in all libraries listed below, for keeping solution viable and without unnecessary costs.

Dashboard makes use of TypeScript - a free and open-source high-level programming language that adds static typing with optional type annotations to Java Script. TypeScript speeds up development experience by catching errors and providing fixes.
Following libraries are being used in dashboard:

- REACT – Redux Toolkit – Redux is a library implementing Flux pattern, for managing and updating application state, using events called "actions". It serves as a centralized store for state that needs to be used across entire application, with rules ensuring that the state can only be updated in a predictable fashion. It allows more efficient control and testing of the application.
- Leaflet - an opensource Java Script library used to build web mapping applications. Leaflet's goal is to be a tool focused on ease of development, focused on efficiency and user-centric design. It seamlessly operates across various platforms, both mobile and desktop, offering comprehensive documentation and extensive plugin options for customization.
- Material UI - an open-source React component library that implements Google's Material Design , Material Design aims to reflect the way objects react to the light and cast shadows in the physical world, operate with bold, intentional elements from print typography and introduce fluidity of movement for their interactions. Material Design uses more grid-based layouts, responsive animations and transitions, padding, and depth effects such as lighting and shadows.
- KeyCloak JS - As an OAuth2, OpenID Connect, and SAML compliant server, Keycloak can secure any application and service as long as the technology stack they are using supports any of these protocols. For some programming languages, Keycloak provides libraries that try to fill the gap for the lack of support of a particular security protocol or to provide a richer and tightly coupled integration with the server.

General SILVANUS Architecture Overview

SAL is the module of SILVANUS responsible for transferring data between different components in SILVANUS. The primary function of SAL is to abstract the object store, offering two key advantages. Firstly, it enables flexibility in managing data at rest, allowing for efficient data management practices. Secondly, it decouples data from user products in a multi-source, multi-client environment, providing support for security, policy, privacy, and business constraints. From the perspective of Dashboard, the ma in purpose of SAL is to store and provide access to data produced by User Products.

Figure 4. showcases the flow of ingestion and provision of data by SAL from external and internal sources, as well as user products. Each ingestion method requires a distinct SAL interface to facilitate communication.

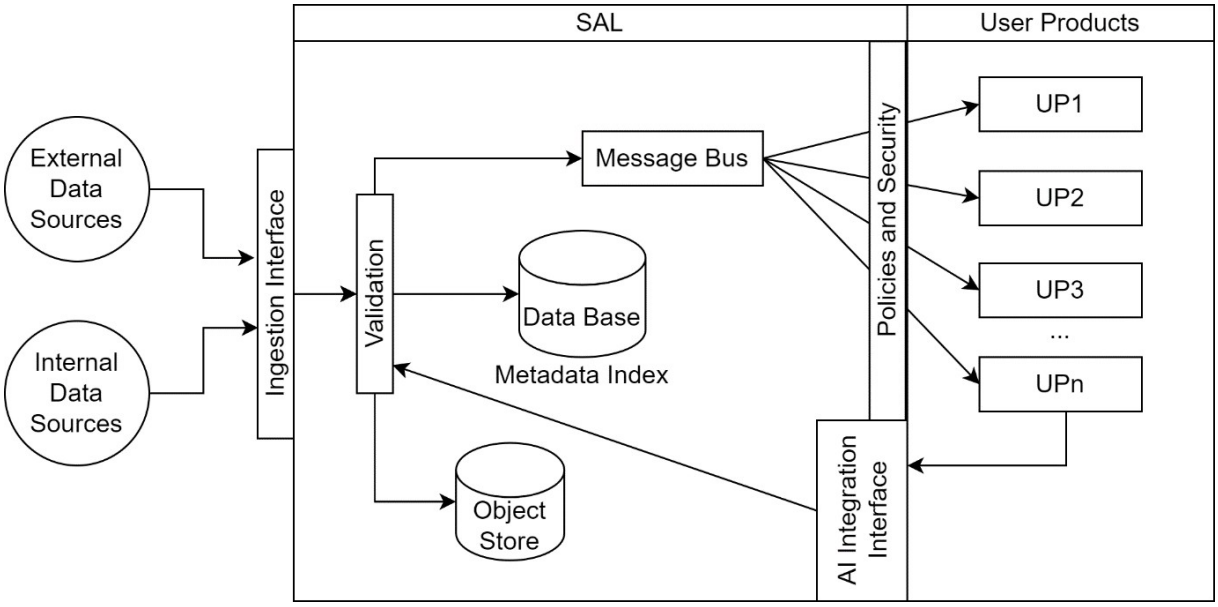


Figure 4. The interaction between SAL and other components in SILVANUS platform

Integration with SAL

The data is produced by the UPs (User Products) and stored in SAL. Between the UI and SAL the middleware exists in the form of Proxy Backend, which is responsible for downloading and transferring the data from SAL to the dashboard. Connection between Proxy Backend and UI is being authorized through Keycloak.

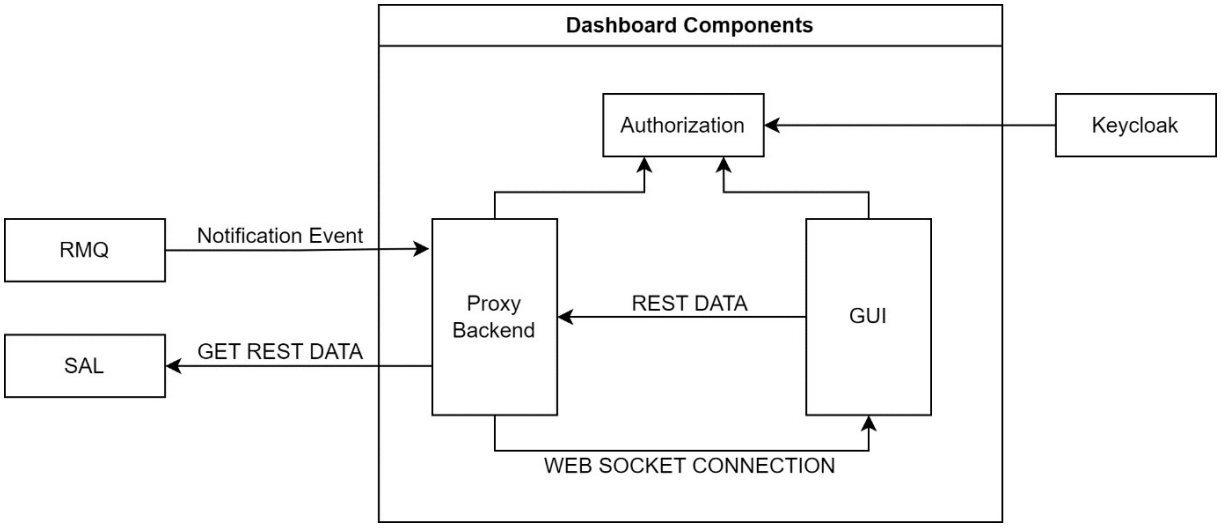


Figure 5. Dashboard – SAL integration general architecture

Proxy Backend – the component performs the task of communicating with the SAL component via a REST API by fetching data and transforming it into the final format that is delivered to UIF. In addition, it receives

notifications from the RabbitMQ queue and forwards them via WebSocket.

Data

There are two different types of data which could be digested by dashboard:

- Geolocalised events (visualised via pins)
- State of the area (visualised via polygon/raster)

Both of those data types are standardized throughout SILVANUS in a certain form of geojson. The data are either static or live, which determines the way in which they are downloaded from SAL. Static data are being downloaded through REST API and live through WebSockets.

Dashboard is able to work with outside GIS servers to manage data flow for less complex layers with static data. The advantage of such a solution is that there is no need to handle data on the side of dashboard and multiple different sources of such data could be applied to the dashboard. On the other hand, there is limited customization and interaction from the dashboard available in such cases.

Dashboard Layers

The SILVANUS dashboard displays a map with several layers representing the work of different User Products (UP). UPs are users' creation in the form of products, content or services. The list with names of UPs and corresponding layers is presented below in Table 3.

Table 3. User Products - Dashboard Layers

UP	Dashboard layer
Fire detection based on social sensing	Social Media Sensing Layer
	ID Social Media Sensing Layer
Fire detection from IoT devices	IoT Layer
Fire detection based on UAV/UGV	UGV Layer
Fire spread forecast	Fire Spread Forecast Layer
Citizens engagement programme and mobile apps	Mobile Subscribers Layer

Depending on the data collected, layers are divided into those displayed in the form of pins and those that represent a certain area. All layers can be enabled simultaneously and each separately. Figure 6 represents a use case diagram of all layers with pins. A pin is a point on the map which geolocation was sent along with other necessary data by a specific device. Currently, there are four layers of this type: Unmanned Ground Vehicle (UGV), Internet of Things (IoT), Social Media Sensing, Mobile Subscribers.

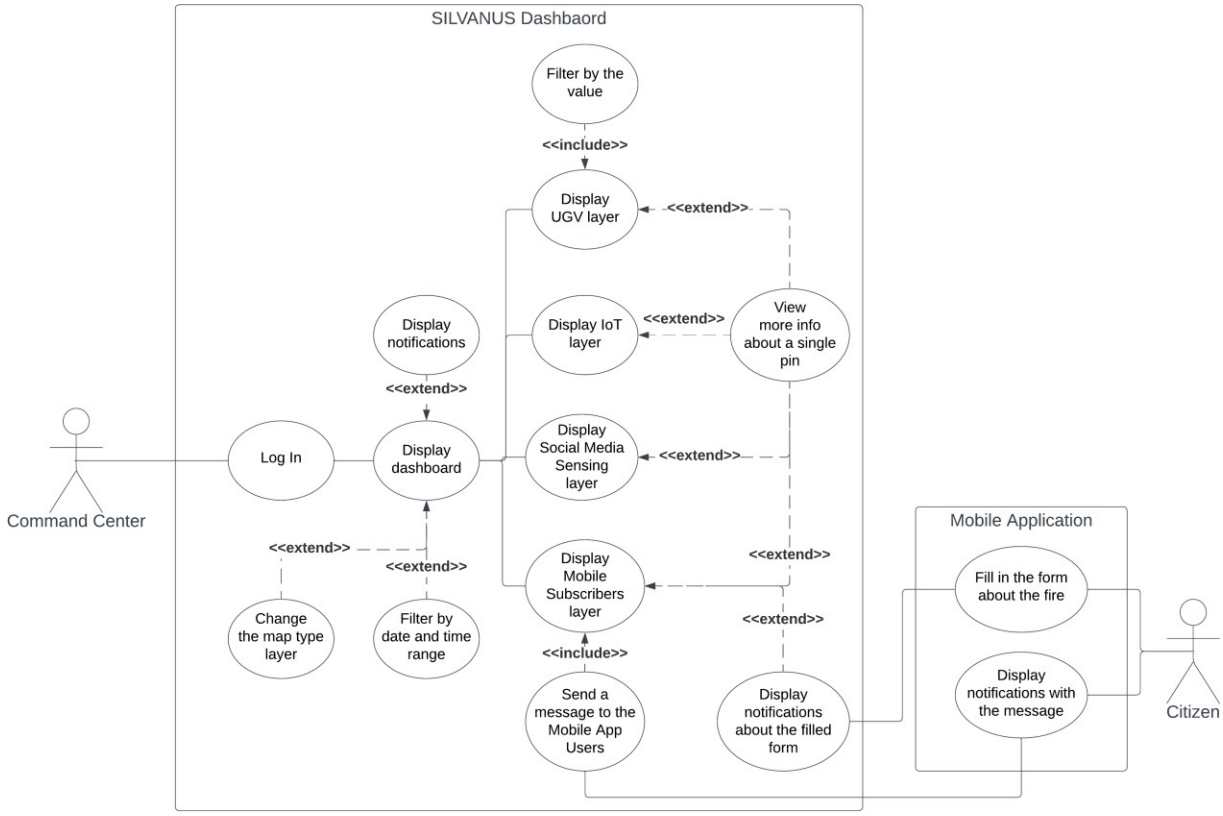


Figure 6. Use Case for SILVANUS Dashboard layers with pins

UGV is a vehicle that operates on the surface of the earth without human intervention on board. These devices move along the previously indicated route, collecting photos and indicators that are sent to the SILVANUS Dashboard every few minutes. They help, among other things, understand the actual state of the field in order to possibly detect a fire. It can be filtered by date, time and value.

IoT is a layer where the device is attached to a fixed and stable place, for example a tree, equipped with sensors and a camera. Similarly, to UGV, it periodically sends data along with photos to the SILVANUS platform. When fire or smoke is detected, it is appropriately recorded with an indicator that affects the color of the pin, thanks to which the dashboard user obtains information about a possible threat faster. The user can filter it by date and time.

Social Media Sensing layer receives alerts from social media, which are generated based on data from social media such as the X platform (formerly Twitter) or Facebook. They contain information about fire, smoke and other fire hazards. It can be filtered by date and time.

Mobile Subscribers is the layer where information comes from citizens who are using mobile devices with the SILVANUS mobile application installed. When they face a fire threat, they can send a photo along with information about it. Moreover, the SILVANUS Dashboard user, having already learned about a fire in a specific area, can send a warning to users of the mobile application. It is possible to filter by date and time.

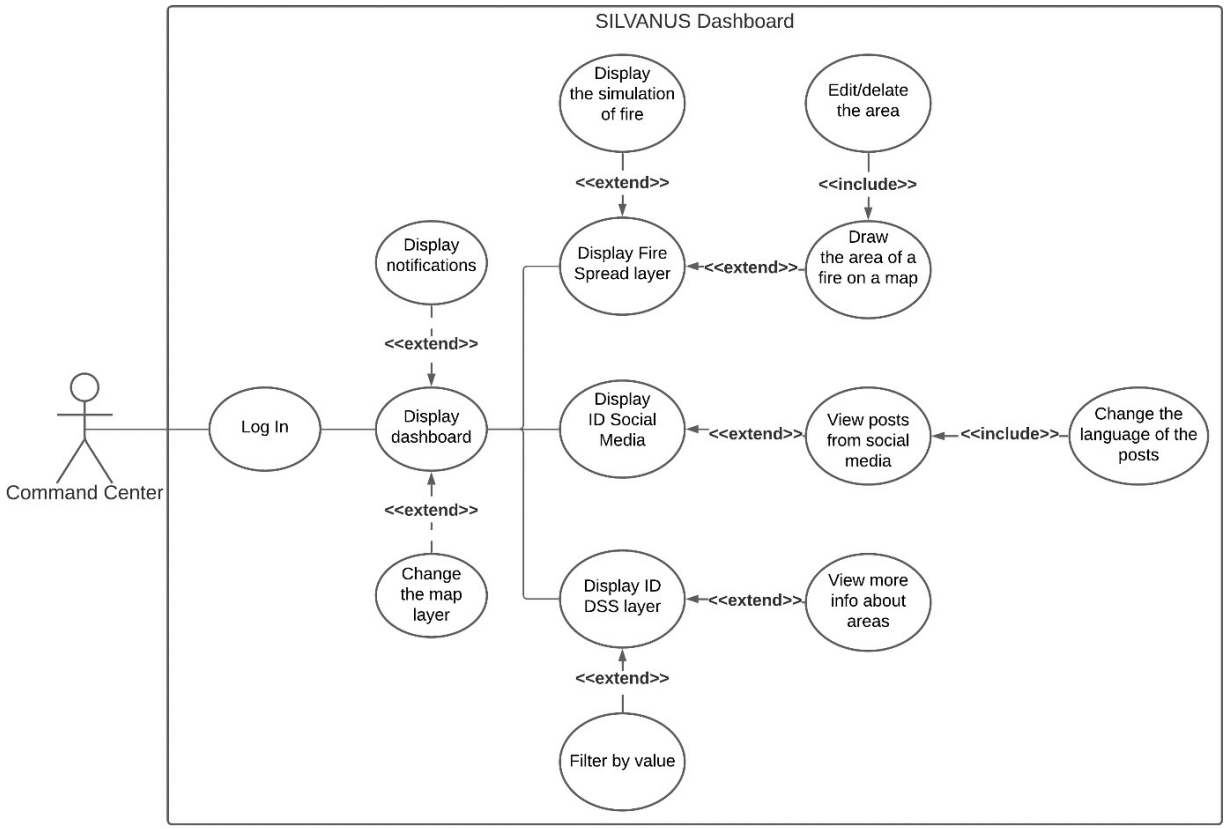


Figure 7. Use Case for SILVANUS Dashboard layers with areas

The next type of layer represents data on areas which is shown on the use case diagram in Figure 7. We distinguish Fire Spread layer (FS), ID Social Media Sensing and Decision Support System for analytics purpose (ID DSS) both prepared for Indonesia region. FS is a layer with simulation of fire spread. The user can view simulation results for a given time in future by moving the slider to a different time range given in minutes. This will display the simulated fire spread, which is shown on the map in the form of a grid. Next, we have layers focusing on the Indonesian area. ID Social Media Sensing layer is very similar to the layer mentioned earlier, they differ primarily in the form of data presentation, in this case they are not pins but areas. ID DSS is the analytical layer that collects statistical data for easy comparison, where the data is presented in the form of charts. Most of the layers can be seen in use on the SILVANUS Dashboard screenshot in Figure 8.

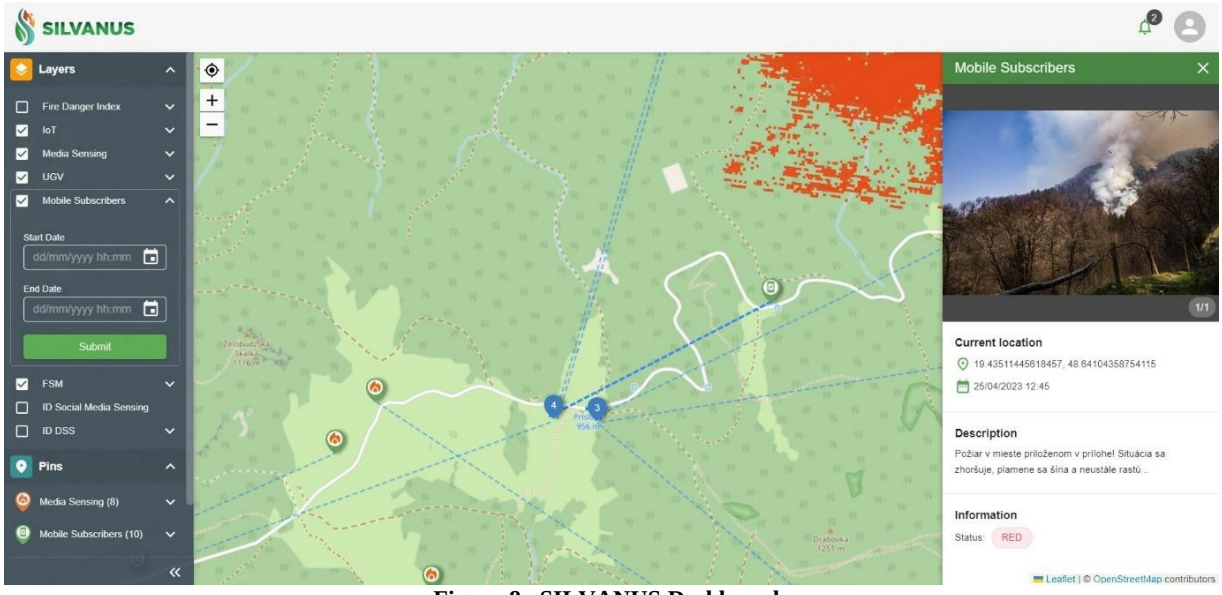


Figure 8. SILVANUS Dashboard

On the left side of the interface there is a panel that allows the user to select the layer with filtering option and the

list of all displayed pins. The main view of the platform is a map which, after selecting a layer, displays the data of interest to the user. Each pin contains information that the user can read by clicking on the pin icon, which will cause the panel to slide out on the right side of the dashboard.

INTEGRATION WITH THE CITIZEN ENGAGEMENT MOBILE APP

The Citizen Engagement Mobile App (CEA) is a specialized mobile application featuring various modules designed to involve citizens in combating forest fires. In addition to citizen engagement modules, such as Guidelines, Quizes, Practical Tips or News, a "Situational awareness and information sharing" module for receiving alerts and sharing event reports in direct communication with firefighters deployed at the scene is also being implemented. The idea is to provide tools for citizens to communicate with firefighters as close to the incidence site as possible. The two main supported use cases are "Fire Warning" and "Fire Reporting". The following Figure depicts the basic actors and use cases.

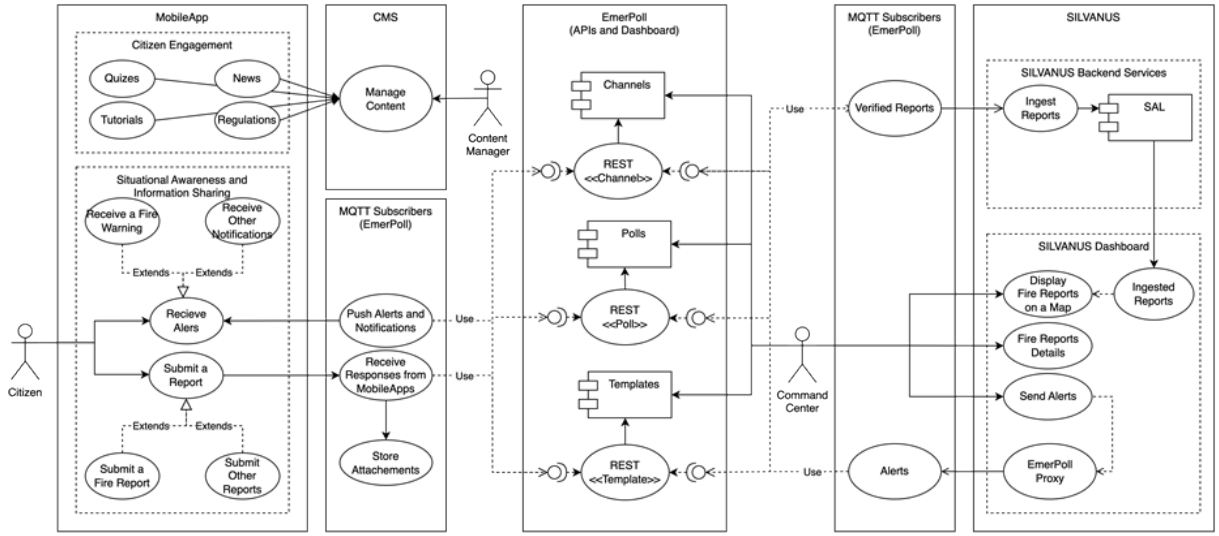


Figure 9. Use Case for the SILVANUS Mobile App and the EmerPoll backend

The main systems involved in the use cases are the mobile application (MobileApp), a content management system (CMS), the SILVANUS platform and the EmerPoll system.

The MobileApp encompasses the "Citizen Engagement" subsystem in which a Content Manager can manage and update the content of the citizen engagement modules and the "Situational Awareness and Sharing" subsystem in which functionalities for receiving alerts and notification as well as for submitting reports about a spotted fire or other events are supported.

EmerPoll is a response collection and aggregation framework which is based on the concept of Polls, Channels and Templates, proposed and described in (Balogh, 2016), for which it exposes both REST APIs for actions as well as a dedicated web-based Dashboard. EmerPoll is the main backend engine making possible the submission, aggregation and validation of individual reports or alerts. MQTT Subscribers are special services which facilitate communication between EmerPoll and other systems by compiling and forwa rding MQTT messages.

The involved SILVANUS system comprises the Backend Services and the SILVANUS Dashboard subsystems. Verified reports coming from the EmerPoll system are ingested using the publish/subscribe paradigm to a storage abstraction layer component (SAL). Subsequently the SILVANUS Dashboard uses such ingested verified reports and displays them on a specialized map layer. Forms are used to send alerts or warnings from the SILVANUS Dashboard via the EmerPoll APIs to the Mobile App users.



Figure 10. Screen of CEA app receiving fire warning region

The CEA application serves as a mobile platform enabling citizens to report sightings of fires (Balogh, 2023) during outdoor activities. In addition to sending the information, users of CEA may receive updates on the severity of fires in their vicinity. Specifically, a polygon marked with a warning icon is displayed on the user's map (shown in Figure 10) to indicate areas of concern. These warning messages are relayed from the EmerPoll service initiated by the SILVANUS Dashboard, where responsible fire-rescue personnel, upon confirming a fire, designate the dangerous area requiring a response. Subsequently, the message is distributed to users through their registered channels as an MQTT message containing a GeoJSON payload with metadata describing the fire incident. In a similar manner, the CEA is also prepared to display evacuation routes from the endangered area, ensuring users have access to critical information for their safety.

CONCLUSIONS AND FUTURE WORKS

Dashboard is a web-based, graphical interface created for SILVANUS platform, which is the main access point for the user to obtain and interact with outcomes of data produced by SILVANUS' various modules. Panel is meant to be used both by representatives from authorities managing fire prevention and regular operators derived from firefighters. Data from different User Products is presented on separate layers in a way which is clear and easy to navigate. Dashboard facilitates timely and easy user access to important information. It provides a variety of tools and parameters upon which the operator is able to determine geolocated occurrences of smoke and fire to put adequate countermeasures to them.

It is important to know that for interface there is no single all-in-a ll solution, or design, which will answer every need from every partner. Each case should be considered separately and be customized based on the individual needs of every partner and characteristics of their region. Final solution will be applied case by case to different regions considering local challenges and accommodate them appropriately.

Dashboard is currently aimed to be visual representation of combined works undergoing in SILVANUS for the partners, authorities and representatives from different regions during pilots and field presentations. In 2023 it was presented during pilots in Podpolanie (Slovakia), Limoges (France), Palangka Raya (Indonesia) and Brisbane (Australia). During those presentations the feedback from possible end users was gathered and is currently being analyzed and considered for further improvements.

Dashboard is constantly evolving upon continuous testing and gathered feedback from actual environments involved in fire prevention, both from perspective of offered functionalities and undergoing improvements in design, usability and user experience. As for 2024 Dashboard is being prepared for upcoming pilots in Czech Republic, Brazil and Italy and the changes identified during previous pilots are being applied.

ACKNOWLEDGMENTS

This work is made under the SILVANUS project which is funded by the EU Horizon 2020 Green Deal program under grant agreement ID: 101037247, Slovak national project iControl APVV-20-0571.

REFERENCES

Balogh, Z., Gatial, E., Hassankhani Dolatabadi, Dlugolinský, Š., Saltarella , M., Scipioni, M. P., Grunwald, D., Przybyszewski, M., Majlingová, A., Brodrechtová, Y., Mojir, K. Y., Olson, N., Levak, J., Chandramouli, K., Communication protocol for using nontraditional information sources between first responders and citizens during wildfires. In Proceedings of the International *ISCRAM Conference*, ISCRAM, 2023, vol. 20, pp. 152-165. ISSN 2411-3387

Battiston, S., Friedemann, M., Gascón, D. M., Viseras, A., Cardíl, A., Mendes, M., Vendrell, J., Barth, B., Nebot, E., & Clandillon, S. (2019). HEIMDALL: a Technological Solution for Wildfires and Multihazard Management Support. *IFBFC2019*. https://heimdall-h2020.eu/wp-content/uploads/2019/03/IFBFC2019_Battiston_Heimdall_accepted.pdf

Few, S. (2013). Information Dashboard Design: Displaying Data for At-a-glance Monitoring.

FOREST EUROPE, 2020: State of Europe's Forests 2020, Ministerial Conference on the Protection of Forests in Europe, 2020

Forzieri, G., Girardello, M., Ceccherini, G., Spinoni, J., Feyen, L., Hartmann, H., Beck, P. S. A., Camps-Valls, G., Chirici, G., Mauri, A., & Cescatti, A. (2021). Emergent vulnerability to climate-driven disturbances in European forests. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-21399-7

Grebner, D. L., Bettinger, P., & Siry, J. (2013). *Introduction to forestry and natural resources. Academic Press.*

Hellmund, T., Manfred Schenk, Schenk, M., Moßgraber, J., Hertweck, P., Reuter, J., & Springer, H. (2022). ELD-BS: the digital situation dashboard for Baden-Württemberg. *ISCRAM*. https://idl.iscram.org/files/tobiashellmund/2022/2413_TobiasHellmund_etal2022.pdf

Hunter, R., & Westerman, G. (2009). *Real business of IT: How CIOs Create and Communicate Value* . Harvard Business Review Press.

Kalabokidis, K., Ager, A. A., Finney, M. A., Athanasis, N., Palaiologou, P., & Vasilakos, C. (2016). AEGIS: a wildfire prevention and management information system. *Natural Hazards and Earth System Sciences*, 16(3), 643–661. https://doi.org/10.5194/nhess-16-643-2016

Mordecai, Y., & Kantsepolsky, B. (2018). Intelligent utilization of dashboards in emergency management. *ISCRAM*. https://dblp.uni-trier.de/db/conf/iscram/iscram2018.html#MordecaiK18

Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., Koutsias, N., Rigolot, É., Barbati, A., Corona, P., Vaz, P. G., Xanthopoulos, G., Mouillot, F., & Bilgili, E. (2011). Landscape – wildfire interactions in southern Europe: Implications for landscape management. *Journal of Environmental Management*, *92* (10), 2389–2402. https://doi.org/10.1016/j.jenvman.2011.06.028

Schubert, B. (n.d.). *Emxsys*. Retrieved April 10, 2024, from http://www.emxsys.com/documentation/default.html

SILVANUS. (2021). *SILVANUS Project*. Retrieved January 23, 2024, from https://silvanus-project.eu/

Technosylva, Inc. (n.d.). *Wildfire Analyst*. Retrieved April 10, 2024, from https://technosylva.com/products/wildfire-analyst/

The Sanborn Map Company, Inc. (n.d.). *Forestry*. Retrieved January 12, 2024, from https://sanborn.com/forestry/



Risk Communication and Community Engagement

Track Chairs

Shelby Knudsen, Trilateral Research

Su Anson, Trilateral Research

Jeannette Annies, University of Stuttgart

Correlation Analysis of the Terrorist Attacks and Social Attributes in The Gulf States

Ting Liu
Joint International Research Laboratory of Catastrophe Simulation and Systemic Risk Governance, Beijing Normal University, Zhuhai 519087, China;
School of National Safety and Emergency Management, Beijing Normal University, Zhuhai 519087, China
202321099011@mail.bnu.edu.cn

Xiaoyong Ni*
Joint International Research Laboratory of Catastrophe Simulation and Systemic Risk Governance, Beijing Normal University, Zhuhai 519087, China;
School of National Safety and Emergency Management, Beijing Normal University, Zhuhai 519087, China
nixiaoyong@bnu.edu.cn

Miaohan Shu
College of Education for the Future, Beijing Normal University, Zhuhai 519087, China
1783963504@qq.com

Dafeng Ma
College of Education for the Future, Beijing Normal University, Zhuhai 519087, China
202211038583@mail.bnu.edu.cn

Jiawei Sun
Faculty of Arts and Sciences, Beijing Normal University, Zhuhai 519087, China
202211079130@mail.bnu.edu.cn

*Correspondence: nixiaoyong@bnu.edu.cn

ABSTRACT

Terrorist attacks are a significant threat to global security as they can cause a large number of deaths, economic losses, and trigger public panic. This paper focuses on the Gulf states as the survey region. This study comprehensively investigates the correlation between national-scale terrorist attacks and various potential factors in the social, political, economic, and military fields, based on the Spearman rank coefficient and gray correlation analysis. The potential factors are divided into two categories: those with a positive correlation and those with a negative correlation to terrorist attacks. The study shows a strong correlation between the adolescent fertility rate (births per 1,000 women aged 15-19) and the occurrence of terrorist attacks. This correlation has not been previously noted in relevant studies, indicating a need for greater attention to be paid to the impact of terrorist attacks on the human rights of children and adolescents.

Keywords

Terrorist attack, Gulf states, characteristic analysis, correlation analysis.

INTRODUCTION

Terrorist attacks involve violent acts against non-military targets, often motivated by political, religious, or ideological goals (Kydd & Walter, 2006; Richards, 2014). The definition used by the Global Terrorism Database (GTD) is adopted in this study, defining terrorist attacks as "the threatened or actual use of illegal force and violence by a non-state actor to achieve political, economic, religious, or social goals through fear, coercion, or intimidation." Terrorist attacks, characterized by obvious violence and inhumanity, can not only cause huge casualties and property losses but also create a panic atmosphere within a certain area for a long time (Holman et

al., 2008; Roy et al., 2011; Yaya, 2009). Moreover, these attacks seriously affect and threaten political security and social stability of all the countries and regions globally, posing a huge threat to international relations and the global economy through ripple effects (Mun et al., 2021). Globally, we have experienced a variety of tragic terrorist attacks. For example, the September 11 attacks were the worst terrorist attacks happened in the United States, resulting in 3,025 deaths and missing. This incident seriously weakened American people's sense of security and formed severe psychological impact on American people who experienced the terrorist attack firsthand (Tong et al., 2016). In 2004, 190 people were killed and more than 1,500 were injured in the Madrid metro bombings (Buesa et al., 2007). On the night of November 13, 2015, a series of terrorist attacks took place in Paris, France, killing 132 people and injuring more than 300, which induced a state of emergency declaration in France. In another well-known case, the multiple bomb attacks in Sri Lanka on April 21, 2019 , resulted in more than 250 deaths and more than 500 injuries. The identification of 42 foreign nationals among the victims caused global shock and widespread concern. Although the GTD indicates a general decrease in the number of terrorist attacks worldwide in recent years, the overall volume remains substantial, with approximately 8,000 attacks recorded in 2020, equating to an average of 22 attacks per day (GTD, 2023). Considering the global impact of terrorist attacks and the serious humanitarian disaster they may bring, research on the characteristics of terrorist attacks itself and the potential related factors are in urgent need.

Terrorist attacks can indeed happen suddenly, but when analyzed on a regional level, patterns and risks of terrorist attacks can often be identified. The academic community has dedicated significant efforts to researching the spatiotemporal evolution and driving mechanisms of terrorism, which can provide valuable insights into the root causes and trends of terrorist attacks. Among the numerous studies, several commonly mentioned factors that are highly correlated with terrorist attacks have been frequently identified, including imperialism, political systems, race and religion, and social factors such as income inequality and unemployment rates (Leite et al., 2019; Sharma, 2013; Tin et al., 2022). Specifically speaking, most of the existing studies engaged in in-depth mining of GTD. Part of the research is the data rearrangement and statistical presentation of the GTD. For example, some scholars searched the GTD for all terrorist attacks against nuclear facilities, nuclear scientists, nuclear transport, and other nuclear industry-related targets in the period from 1970-2020, and found that terrorist attacks carried out by non-state perpetrators were rare, and none of the attacks finally resulted in radioactive fallout or environmental contamination (De Cauwer et al., 2023). Part of the research is knowledge mining or knowledge discovery based on GTD datasets. For example, in terms of temporal and spatial characteristics, a study discovered countries located in a hot spot neighborhood of terrorist attacks are likely to experience a large increase in the number of terrorist attacks in the next period. This suggests that the spatial distribution of terrorist attacks is an important feature, and indicates that terrorist attacks have a characteristic of spatial transmission (Braithwaite et al., 2007). In terms of influencing factors, it has been suggested by some studies that states that fail to respond to internal political strife, ethnic conflict, or "stateless areas" are more susceptible to terrorist attacks (Piazza, 2007). This highlights the importance of governance in a country or region when it comes to preventing terrorism. Of course, some other factors like religiosity (Varenia et al., 2021), the ethnicism (Cho, 2010), democracy and civil conflicts (Kim et al., 2020; Morris et al., 2021), the popularity of the national leaders (Krueger et al., 2009) were also revealed to be highly related to terrorist attacks.

As to the terrorist attacks themselves, GTD provides detailed information on the location, type of attack, weapons used, and target type of terrorist incidents. Researchers have employed hierarchical analysis and K-means clustering analysis to categorize these terrorist attacks (Li et al., 2021). Additionally, studies have evaluated the heterogeneity of attack types worldwide using methods such as chi-square tests (De Cauwer et al., 2022; Ulmer et al., 2022). For instance, Pan conducted a quantitative analysis, fully exploring the correlations between terrorist organizations and other event attributes in the GTD. Based on machine learning techniques, he proposed a method to predict terrorist organizations using other indicators (Pan, 2021). Last but not least, some studies aim to provide additional information beyond the GTD dataset. For instance, an evaluation index system for the harmfulness of terrorist attacks could be constructed using the projection pursuit model. This system could be used to score all terrorist attacks in the GTD based on their consequences. This information wa s a valuable addition to the GTD dataset. Time series methods, such as the grey metabolic model (Deng et al., 2021), can be used to predict the future risk of terrorist attacks. These studies enriched and deepened our understanding of the GTD dataset. However, the factors linked to terrorist attacks usually intertwine to form a complex network. It's sometimes hard to distinguish which factors are the causes of terrorist attacks and which are the effect of terrorist attacks, since some factors are mutually causal with terrorist attacks, which is also a point worth further study.

The Middle East is a region with a high-incidence of terrorism. According to the GTD, the Middle East has an average of 1,165 terrorist attacks per year from 1970 to 2020, accounting for 27% of the global total. The Gulf region has long been the focus of geopolitics and conflicts, which has created a breeding ground for terrorism and seriously affected the security and stability of local countries and regions. From 2014 to 2021, the Islamic State (ISIS) carried out several terrorist attacks in the region, including violence in areas controlled by Iraq and Syria, as well as terrorist attacks in other countries. This has resulted in a severe security situation in relevant areas.

Multiple factors, including political conflicts, wars, terrorism, social inequality, and natural disasters, have continuously caused unrest in the Middle East, leading to humanitarian disasters and crises in many countries. The issue is also highlighted by the adolescent fertility rate indicator studied in this research. According to the World Bank data, the countries and regions with the highest rankings are mainly concentrated in Africa and the Middle East. Taking Iraq as an example, the adolescent fertility rate in 2021 reached 62.19‰, which although has decreased compared to previous years, is still significantly higher than the world average of 42 ‰. As a foundation for subsequent global-scale studies, this paper selects the Gulf states as the survey region, including Bahrain, Iran, Iraq, Kuwait, Qatar, Saudi Arabia, the United Arab Emirates and Oman. Spearman correlation analysis and Gray correlation analysis were used to explore the characteristics of terrorist attacks and related factors. It should be noted that correlation analysis in this paper is distinct from causation analysis, which means that causality cannot be strictly inferred. The causation analysis of terrorist attacks is relatively more complex and will be furtherly studied in the future. However, the authors believe that the paper could still have a great contribution. To the best of the authors' knowledge, this paper makes the first attempt to reveal the correlation of social multi-source data and terrorist attacks in the Gulf states. Also, some of the key findings provided by this paper are original and represent the first proposals.

OVERALL STRUCTRUE

As a preliminary study in the authors' research plan, the overall framework of this paper is straightforward. In this study, the Gulf states are selected as the survey region. Two main datasets, GTD and the World Bank dataset, are utilized in this study. The basic characteristics of terrorist attacks are analyzed in depth using the GTD, while additional associated factors are retrieved from the World Bank dataset. The Spearman's rank coefficient analysis and Grey relation analysis are two well-known and commonly used methods. They are employed in this study to investigate the correlations between terrorist attacks and related factors. Several important findings could be obtained by this study by quantitatively defining the relationship between terrorist attacks and the factors of socio-economic system, such as identifying which factors are most relevant to terrorist attacks. The results from this study are crucial for addressing the risk of terrorist attacks and assisting the communities in making targeted preparations.

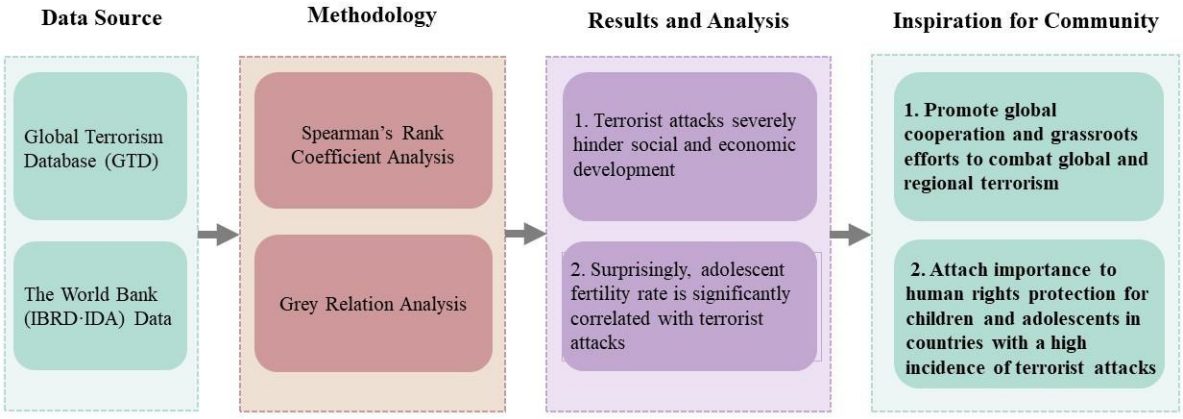


Figure 1. The overall structure of this paper

STUDY AREA AND DATASET

Survey Region

The Persian Gulf, abbreviated as "the Gulf", is one of the most important inland seas in the world, located in southwestern Asia between the Arabian Peninsula and Iran. The term "Gulf States" is a widely accepted concept, rather than a strictly defined geographic term with clear boundaries. In this study, the "Gulf States" refer to the eight countries along the Persian Gulf coast: Bahrain, Iran, Iraq, Kuwait, Qatar, Saudi Arabia, United Arab Emirates and Oman (Figure 2). The profiles of the eight countries were collected and briefly organized (Table 1), comprising information on area, population, religion, language, political system, economy, as well as military affairs.

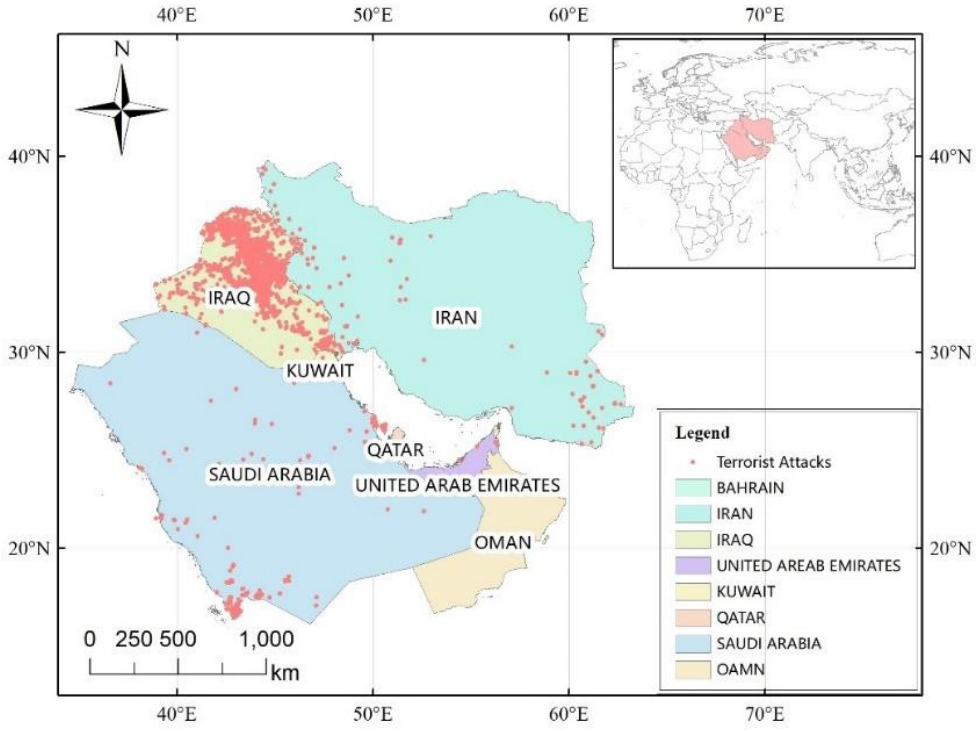


Figure 2. Terrorist Incidents Based on Geographic Location of Gulf States in 2001-2020

Table 1. The basic situation of eight Gulf states

Country Name	United Arab Emirates	Oman	Bahrain	Kuwait	Iraq	Iran	Qatar	Saudi Arabia
Area	83,600 sq.km	309,500 sq.km	780 sq.km	17,818 sq.km	438,300 sq.km	1,645,000 sq.km	11,521 sq.km	2,250,000 sq.km
Population	10.17 million	5.08 million	1.5 million	4.824 million	43.35 million	85.02 million	2.97 million	32.18 million
Religion	Mostly Islamic	Islam is state religion	Mostly Islamic	Islam is state religion	Mostly Islamic	Islam is state religion	Mostly Islamic	Islam is state religion
Language	Arabic; General English	Arabic; General English	Arabic; General English	Arabic	Arabic; Kurdish	Persian	Arabic; General English	Arabic
Political System	The highest authority: The Federal Supreme Council	Hereditary monarchies; Political parties are prohibited	Hereditary monarchies; Political parties are prohibited	Hereditary monarchical emirates	Federalism; More than 200 political parties and entities	Theocratic political system; Presidential cabinet system	Hereditary monarchies; Political parties are prohibited	Hereditary monarchies; Political parties are prohibited
Economy in 2022	GDP per capita: $47,700,000; GDP growth rate: 8.9%	GDP per capita: $19,000; GDP growth rate: 4.3%	GDP per capita: $23,500,000; GDP growth rate: 4.9%	GDP per capita: $39,000; GDP growth rate: 7.6%	GDP per capita: $5,021; GDP growth rate: 8.3%	GDP per capita: $4,388; GDP growth rate: 4%	GDP per capita: $81,800,000; GDP growth rate: 4%	GDP per capita: $34,400,000; GDP growth rate: 8.7%
Military Affairs	Voluntary military service system; Total military strength: 56,500	Compulsory military service system; Total military strength: 43,500	Voluntary military service system; Total military strength: 11,800	Compulsory military service system; Total military strength: 23,000	Recruitment system; Total military strength: 778,000	Compulsory military service system; Total military strength: 1.27 million	Voluntary military service system; Total military strength: 12,000	Voluntary military service in peacetime and compulsory military service in wartime; Total military strength: 230,000

Dataset

The datasets utilized in this study include the GTD, which covers records of terrorist attacks in the Gulf States region from 2001 to 2020 (the latest available data is updated to June 2021). GTD is an open-source database operated by the National Center for Terrorism at the University of Maryland and the Consortium for the Study and Research of Terrorism. It is widely recognized as the most comprehensive and sophisticated database on terrorism worldwide. The database incorporates data from reports published by authoritative media organizations. It includes information on the time and location of terrorist attacks, targets, attackers, and the number of people killed, injured and so on. Statistics show that there were 28,264 attacks in the Gulf States from 2001 to 2020. To ensure accuracy, data records with incomplete latitude and longitude information were excluded, ultimately 28,091 terrorist attacks in the Gulf states region were involved in this study.

This study selects eight socio-economic indicators as related factors (Table 2), based on the relevant studies mentioned in the introduction section and using the World Bank database as a reference. These indicators are:

- Per capita GDP (constant 2015 dollars);
- Unemployment rate (the ratio of the total number of unemployed to the total number of labor force);
- Net migration (net migration is the total number of net migrants during the period, i.e., the total number of migrants who migrated abroad minus the number of migrants who migrated into the country each year, including both permanent and non-permanent residents. The number of net migrants in that year is then divided by the country's population in that year);
- Adolescent fertility rate (the number of births per 1,000 15~19-year-old females);
- Military strength (armed forces personnel as a percentage of the total labor force);
- Urban population (the percentage of the urban population compared to the total population);
- Agricultural land (per capita agricultural land refers to land used for short-term cropping, short-term pasture for mowing or grazing, vegetable gardens for supplying the market, and gardens for own use, as defined by the Food and Agriculture Organization of the United Nations (FAO). Recreational land resulting from the conversion of farming practices is not included);
- Commodity trade (as a percentage of GDP).

In addition, considering the differences in each country's population, economy, etc., the Terrorist incident rate is defined as the number of terrorist attacks in a given year divided by the population of that country in the same year.

Table 2. Basic information on related socio-economic indicators to terrorist attacks

Index	Socio-economic variable	Abbreviation
x_1	Per capita GDP ($)	PCG
x_2	Unemployment rate (%)	UR
x_3	Net migration (%)	NM
x_4	Adolescent fertility rate (persons/1000)	AFR
x_5	Military strength (%)	MS
x_6	Urban population (%)	UP
x_7	Agricultural land (hectares)	AL
x_8	Commodity trade (%)	CT
y	Terrorist incident rate (cases/million people)	TIR

METHODOLOGY

The Correlation Analysis Method

Correlation analysis is a straightforward and accessible analytical method for measuring how quantitative data are related to each other, and it is commonly used to determine the degree of correlation between variables. To

calculate the strength of correlation between two different factors, a correlation coefficient is introduced below, that is the Spearman Correlation Coefficient (R_S), as shown in Eq. (1). The Spearman's Rank Correlation Coefficient (R_S) offers greater flexibility compared to the Pearson Correlation Coefficient (R_p), which is typically limited to describing linear relationships assuming a normal distribution. R_S is often used to assess the ranking of relationships, particularly in scenarios where variables possess ordinal properties, do not follow a normal distribution, or when the distribution type is unknown. The flexibility arises from its less restrictive data requirements. Current research often focuses on investigating ordinal relationships between variables, specifically examining whether an increase in one variable is accompanied by an increase in another. Therefore, the formula for calculating Spearman's correlation coefficient is chosen based on research needs. To calculate the Spearman's rank correlation coefficient (R_S), define two data sequences, $\{x_i\}$ and $\{y_i\}$, of equal length, where n is the number of data points in each sequence. Assign rank values (denoted as $\{p_i\}$ and $\{q_i\}$) to each data point in $\{x_i\}$ and $\{y_i\}$ respectively, based on their position from smallest to largest. Then, for each corresponding data point in the two sequences, calculate the difference $\{d_i\}$ between their rank values. Note that variables with the same value should be assigned the same rank, and in case of ties, the average rank is commonly used. The variables $\bar{p}$ and $\bar{q}$ are denoted as the mean ranks, respectively.

$$R_s = \frac{\sum_{i=1}^{n}(p_i - \bar{p})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^{n}(p_i - \bar{p})^2}\sqrt{\sum_{i=1}^{n}(q_i - \bar{q})^2}} = 1 - \frac{6 * \sum_{i=1}^{n} d_i^2}{n\,(n^2 - 1)} \quad (1)$$

R_s is an index that quantifies the strength of correlation in the range of [-1, 1]. The closer the value of R_s is to either 1 or -1, the stronger the correlation. R_s >0 indicates a positive correlation, while R_s <0 indicates a negative correlation.

When performing correlation analysis, significance level tests are routinely conducted to assess whether the observed correlation is statistically meaningful. This means that the correlation is unlikely to be due to random factors. In statistics, the p-value is used to assess the probability of the sample data appearing under the null hypothesis, thus testing the statistical significance of an effect. The Spearman's rank correlation coefficient, for example, can be used to measure the strength and direction of the relationship between two variables, and the resulting correlation coefficient can then be used in a hypothesis test to calculate a p -value. Generally speaking, the smaller the p-value, the more significant the correlation is, indicating that the observed effect is less likely to be due to random error. To visually communicate the significance of these results, asterisks (*) are often used as markers. Specifically, a single asterisk (*) is used when the p-value for the correlation is less than 0.05, indicating statistically significant correlation. Two asterisks (**) are used for p-values less than 0.01, emphasizing the reliability of the correlation. Three asterisks (***) are reserved for p-values less than 0.001, which usually serve as highly convincing evidence of a strong and reliable correlation between the two variables.

The Grey Relation Analysis Method

Grey Relation Analysis (GRA) is a statistical method for multifactor analysis. It determines the degree of closeness between the series of the reference data and the series of several comparison data, by comparing the geometric similarity of their curves and calculating the degree of correlation between the curves. The more similar the curves are, the greater the correlation between the corresponding sequences will be, and vice versa. The reference sequence is usually the sequence that represents the overall development trend of the system, while the comparison sequence requires analysis. Set $x_0' = \{x_0'(k) | k = 1, 2, 3, \ldots, n\}$ as the reference series, $x_i' = \{x_i'(k) | k = 1, 2, 3, \ldots, n\}$ $(i = 1, 2, 3, \ldots, m)$ as the comparison series. As the factors hold varying physical significance, the series data may not always share the same order of magnitude. This can impede comparison and make it difficult to draw accurate conclusions. To address this, it is necessary to convert all sequences to the same order of magnitude. This is typically achieved by dividing each element in the sequence by the average (as shown in Eq. (2)) or first element of the sequence (as shown in Eq. (3)). The grey correlation coefficient between the reference sequence and the comparison sequence is calculated using Eq. (4). Here, $x_i(k)$ denotes the k-th element of the i-th sequence, and $x_0(k)$ represents the k-th element of the reference sequence. The grey correlation degree of the i-th sequence r_{0i} is the average value of all grey correlation coefficients between it and the reference sequence (as shown in Eq. (5)). The rank of the i-th sequence is determined by comparing its grey correlation degree with other sequences. The resolution coefficient, ρ, should be determined based on the specific background and requirements and is in the range of (0,1). The default value of ρ is always set to be 0.5. r_{0i} denotes the gray correlation of x_0 with x_i. The closer to 1 the value of r_{0i} is, the better the correlation is.

$$x_i(k) = \frac{x_i'(k)}{\frac{1}{m}\sum_{k=1}^{m} x_i'(k)} \quad (2)$$

$$x_i(k) = \frac{x_i'(k)}{x_i'(1)} \tag{3}$$

$$\xi_i(k) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|} \tag{4}$$

$$r_{0i} = \frac{1}{m} \sum_{k=1}^{m} \xi_i(k) \tag{5}$$

RESULTS

Time Evolution Characteristics of Terrorist Attacks

Between 2001 and 2020, the eight Gulf states experienced a total of 28,091 terrorist attacks, averaging 1,405 attacks per year. The number of deaths totaled 81,979, averaging 7,037 deaths per year, and the number of injuries totaled 140,744, averaging 4,099 injuries per year. Figure 3 illustrates the temporal evolution of terrorist attacks in all the eight Gulf states. Since the 21st century, the terrorist situation in the Gulf countries has shown significant fluctuations. The number of terrorist attacks, deaths, and injuries has exhibited a trend of growth followed by a decrease, reaching its peak in 2014. These trends are relatively consistent, with two significant turning points in 2007 and 2014. Fatalities resulting from terrorist attacks in 2020 decreased by 36.63% from the previous year to 531, marking the fourth consecutive year of decline since 2016. This decline is closely linked to an improvement in the intense conflict in Iraq.

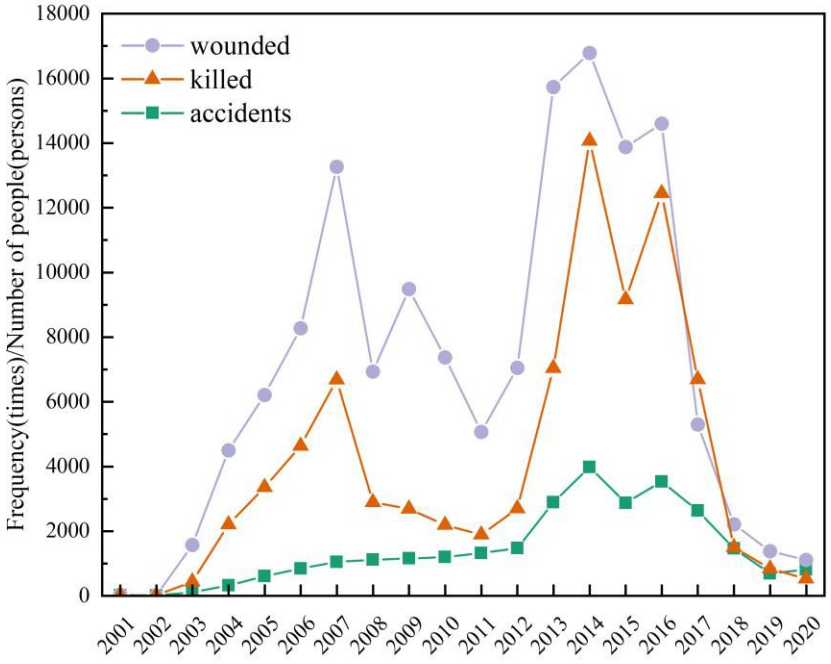


Figure 3. Interdecadal characteristics of terrorist attacks in the eight Gulf states

Out of the 28,091 terrorist attacks, the number of casualties from 841 terrorist attacks was unrecorded by GTD and therefore unknown. The remaining 27,250 terrorist attacks were further analyzed. Figure 4 shows the distribution of deaths in these attacks. Of all incidents, 64.39% resulted in fatalities in 17,545 cases, indicating that terrorist attacks are a considerable cause of casualties, despite the fact that the majority of terrorist attacks result in only a few fatalities.

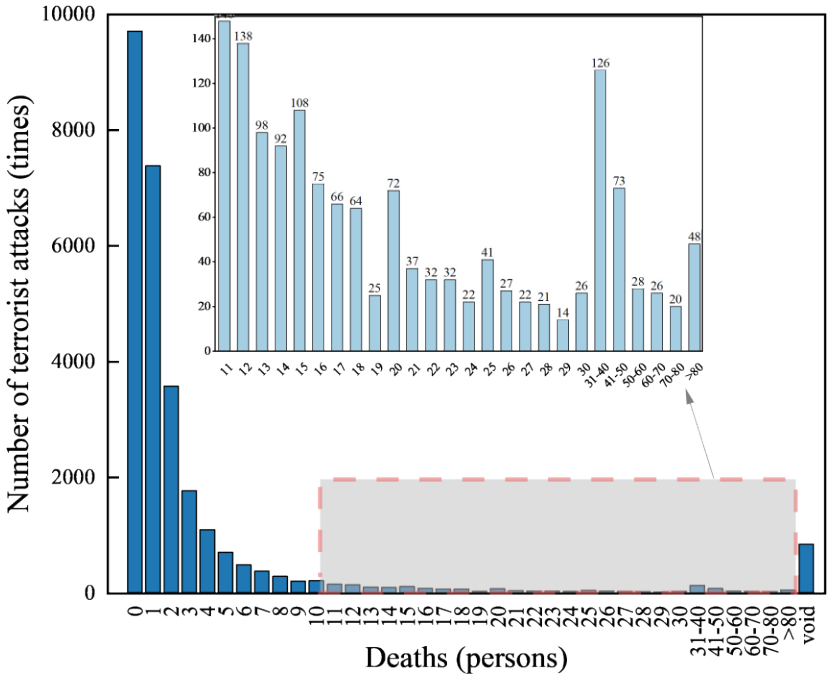
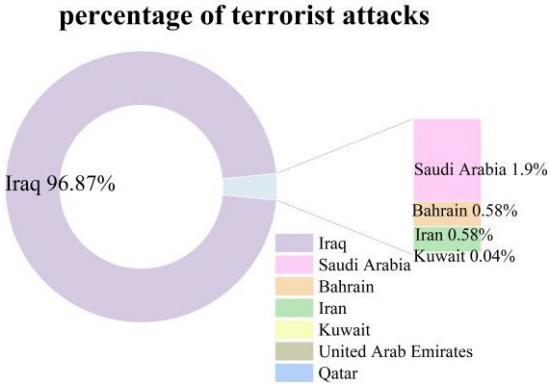


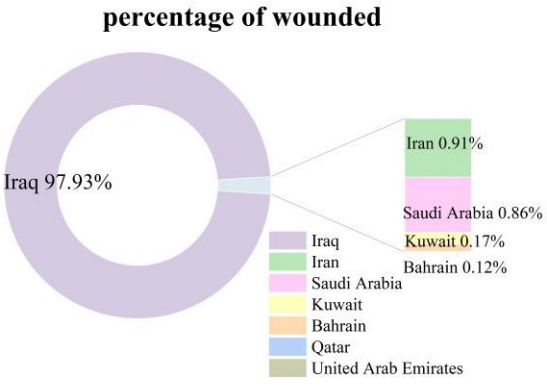
Figure 4. Characteristic analysis of casualties resulting from different terrorist attacks in the eight Gulf states

Regional Distribution Characteristics of Terrorist Attacks

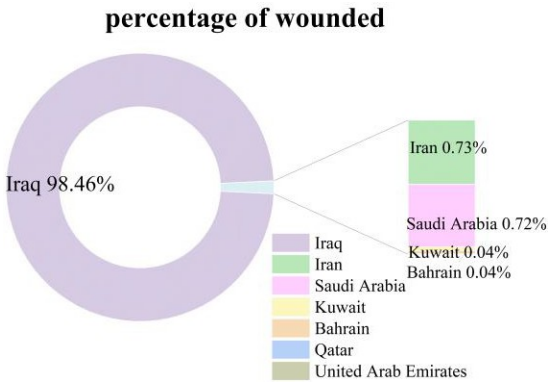
This study examines the geographical distribution of terrorist attacks in the Gulf states between 2001 and 2020, using latitude and longitude data from GTD. The analysis reveals that the majority of terrorist incidents in the Gulf states during the 21st century occurred in Iraq and Iran. The number of deaths in Iraq has reached 80,719, accounting for 98.46% of the total. Additionally, 137,830 people have been injured, accounting for 97.93%. Furthermore, there have been 27,212 terrorist attacks, accounting for 96.87% (Figure 5). These statistics confirm that Iraq is a high-risk area for terrorist incidents in the Gulf region. This phenomenon may be related to the war in Iraq that began in 2003 and the establishment of the terrorist organization I SIS in 2014.



(a) Quantitative characteristics of terrorist attacks in different countries



(b) Quantitative characteristics of the injuries in terrorist attacks in different countries



(c) Quantitative characteristics of the casualties in terrorist attacks in different countries

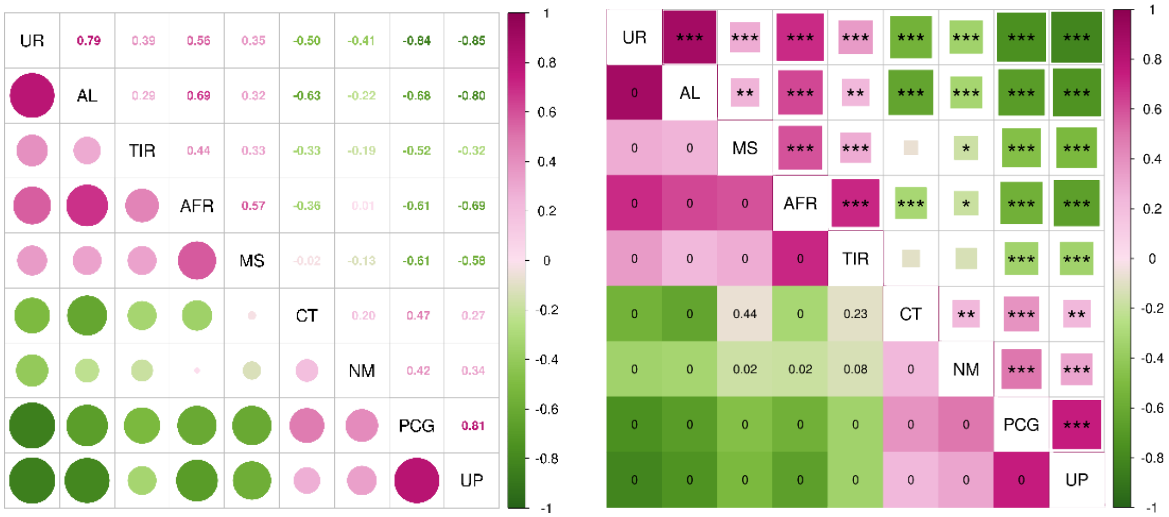
Figure 5. Regional distribution characteristics of terrorist attacks in the eight Gulf states

Correlation Analysis

Terrorism often arises from confluence of factors, such as social instability, political conflict, economic hardship and religious, racial or ethnic differences. These factors may lead to an increase in social tensions, social discontent, and extremist thinking, thus providing a breeding ground for terrorism. At the same time, the occurrence of terrorism will further aggravate social unrest and trigger a series of political, economic and livelihood problems. It can be seen that the terrorist attacks and socio-economic factors are often causal and interdependent, which is quite complex and not yet be confirmed in this field. In this paper, the causal ana lysis of terrorist attacks has been temporarily shelved in favor of a mainly correlational analysis. From the perspective of direct correlation, we have analyzed which factors might be highly correlated with terrorist attacks, and a more in-depth analysis is planned for the future. A Spearman correlation analysis of terrorist attacks in the Gulf States from 2001 to 2020 explores the relationships between the various elements listed in Table 2 above, as shown in Figure 6.

Upon careful observation of the color legend on the right side of Figure 6(a), a clear visual correlation between color and the intensity of correlation can be established. A gradual shift in color from dark to light red corresponds to a decrease in the numerical value from near 1 to 0, indicating a gradual weakening of the positive correlation. Conversely, the transition from deep green to lighter shades corresponds to a decrease in the numerical value from near -1 to 0, signifying a gradual weakening of negative correlation. The figure displays the correlation coefficients between the variables. The right triangular matrix displays the correlation coefficients among various variables, providing precise numerical values regarding their degrees of associatio n. The left triangular part employs the size of the circle as a visual aid, where a larger circle area represents stronger correlation, thus enhancing the visualization of correlation strength. Simultaneously, the circle's color aligns with the legend, emphasizing the correlation's positive and negative directions. Figure 6(b) displays the correlation analysis's significance level. The combination of p-values and asterisks (*) enables quick identification of statistically significant correlations. A value of 0 on the graph corresponds to a p-value less than 0.01, indicating high statistical significance. Similar to Figure 6(a), the colors, size, and depth of the graphs continue to indicate the strength of the correlations. In this figure, they are also integrated with the significance level, providing a more comprehensive analytical perspective.

Figure 6(a) demonstrates a positive correlation between TIR and the influencing factors of UR, AL, AFR, and MS, while a negative correlation exists between TIR and CT, NM, PCG, and UP. Figure 6(b) reveals that the p - values for the correlations between terrorist attacks and the socio-economic factors are less than 0.01, with the exception of NM. Consequently, the majority of the socio-economic factors, excluding net migration, exhibit significant relationships with terrorist attacks based on the presented da ta.



(a) Spearman correlation coefficient (b) Significance level and p-value

Figure 6. Correlation analysis between terrorist attacks and different socio-economic factors

Correlation between TIR and PCG, CT, UR

Specifically, the frequency of terrorist incidents exhibits a significant negative correlation with per capita GDP (PCG) and commercial trade as a percentage of GDP (CT), with correlation coefficients of -0.52 and -0.33, respectively. While a significant positive correlation with the unemployment rate (UR) was observed, with 0.39. It is widely acknowledged that a decrease in the number of terrorist incidents in a region is associated with increased stability and economic development, and vice versa. GDP per capita commonly serves as a key measure of a region's economic prosperity. A high GDP per capita indicates that, on average, people in the region enjoy greater economic resources and benefits, leading to improved quality of life and a more stable social environment. Conversely, lower GDP per capita generally implies economic hardship and poverty, which may lead to an increase in social tensions and discontent, subsequently elevating the risk of terrorist incidents. Therefore, it is not difficult to understand that GDP per capita (PCG) could serve as a reliable index to characterize the intensity of terrorist attacks. Similarly, the healthy development of commercial trade is usually associated with economic dynamism and international cooperation. A vibrant and thriving business environment encourages increased employment opportunities, raises income levels, and stimulates economic development. In contrast, regions with underdeveloped or restricted trade are often at risk of economic decline and social instability . High unemployment, severe poverty, and economic underdevelopment may lead to conflict and provide fertile ground for terrorism.

Correlation between TIR and UP, AL

The correlation coefficient value of -0.32 indicates a highly significant negative correlation between the frequency of terrorist attacks and the ratio of urban population to total population (UP). Additionally, there is an extremely significant positive correlation of 0.29 between the frequency of terrorist attacks and the number of hectares of arable land per capita (AL). This result implies that higher urbanization tends to lead to a lower incidence of terrorist attacks. It is well known that the UP ratio and AL per capita are important indicators of urbanization, which is usually accompanied by economic development, social progress, and improved infrastructure. On the one hand, urban areas usually offer a higher standard of living and greater employment opportunities than rural areas, potentially reducing economic pressures and social grievances that can contribute to the risk of terrorist incidents. Moreover, urban areas tend to have more social communication, promoting interaction among people and strengthening social ties and community participation. These stronger social ties help to build harmonious social relations and increase social stability, thus reducing the nurturing environment for terrorism. On the other hand, terrorist attacks can significantly impede the normal development and urbanization of cities, including a country's overall urbanization level. In conclusion, this outcome aligns with people's perceptions.

Correlation between TIR and MS

A correlation coefficient value of 0.33 indicates a significant positive correlation between the number of terrorist attacks and military strength (MS), measured as the proportion of active military members in the labor force.

According to World Bank data, the countries with the highest percentage of armed forces personnel in their labor force from 2001 to 2020 are predominantly located in Asia and Africa, including Eritrea, the Democratic People's Republic of Korea, and Jordan. It is worth noting that most of these countries are developing nations with limited economic growth. Their economic development primarily rely on agriculture, industry, or resource exploitation. They often exhibit geopolitical complexity, with border disputes or conflicts with neighboring countries. As a result, these countries tend to have stronger militaries, which does not necessarily translate into better control of terrorism. In fact, the outlook for terrorism in these countries is not optimistic. On the contrary, a stronger military sometimes leads to greater instability in the country, increasing the risk of terrorism.

Correlation between TIR and AFR

Frequency of terrorist attacks shows an extremely significant positive correlation with the adolescent fertility rate (AFR), denoted by the number of births per 1,000 females aged 15 -19, with a value of the correlation coefficient of 0.44. AFR varies across countries and regions, and is influenced by various factors, such as sociocultural attitudes, economic conditions, education level, and access to sex education and contraception. Based on 2021 World Bank data, the adolescent fertility indicator ranks African countries such as Niger, Mozambique, and Somalia at the top. Notably, these countries also face significant challenges from terrorism. On one hand, the frequent occurrence of terrorist incidents may trigger social unrest and affect the normal life and studies of adolescents, thus impacting their academic performance and future development. On the other hand, an increase in the fertility rate during adolescence may lead to greater socio-economic pressures, including education and management of troubled youth, and youth social participation. The risk of terrorist attacks may increase due to extremist religious thoughts that can arise from low levels of education, neglect of duty, and engaging in inappropriate business during adolescence.

Grey Relation Analysis

Table 3 illustrates a close and strong correlation between TIR and multiple socio-economic indicators through the application of gray correlation analysis. These findings align with the previous correlation analysis. The top three most significant factors are the adolescent fertility rate (AFR), agricultural land per capita (AL), and unemployment rate (UR), exhibiting gray correlation coefficients of 0.9013, 0.8913, and 0.8892, respectively. Notably, the relationship between terrorist attacks and the AFR is particularly strong, highlighting AFR as the most highly correlated factor.

Table 3. Gray correlation analysis between TIR and different factors

Factor	Grey correlation coefficient	Rank
AFR	0.9013	1
AL	0.8913	2
UR	0.8892	3
MS	0.8876	4
CT	0.8720	5
PCG	0.8621	6
UP	0.8693	7
NM	0.8521	8

Comparing the data on adolescent fertility rates in the eight Gulf region countries (Figure 7), it is clear that Iraq has a significantly higher fertility rate than the other countries. Given the significant number of terrorist attacks that have taken place in Iraq between 2001 and 2020, there is a strong likelihood of a correlation between these attacks and the fertility of adolescents. This underscores the urgency of prioritizing educational opportunities, well-being, and human rights for adolescents in terrorist incident-prone areas. Furthermore, in Iraq, the literacy rate of young women aged 15-24 was below the world average until 2016, indicating that their right to education for this group is not fully guaranteed. Additionally, World Bank data reveals that in 2004, there were only 0.006 community health workers per 1,000 people in Iraq, equivalent to just 6 community health workers per 1,000,000 people. This statistics highlights the extreme inadequacy of community health services, which could pose a serious threat to the health and well-being of children and adolescents. From a global perspective, it is noteworthy that in

certain countries or regions, such as Gaza, Kenya, and Haiti, adolescents and their families face life-threatening crises due to political unrest, civil strife, violence, extreme poverty, and natural disasters. These crises encompass displacement, infectious diseases, and malnutrition. In contrast, the deadly threat posed by terrorist attacks on adolescents has not received sufficient attention. The paper suggests a potential association between terrorist attacks and the human rights of adolescents, which deserves further exploration.

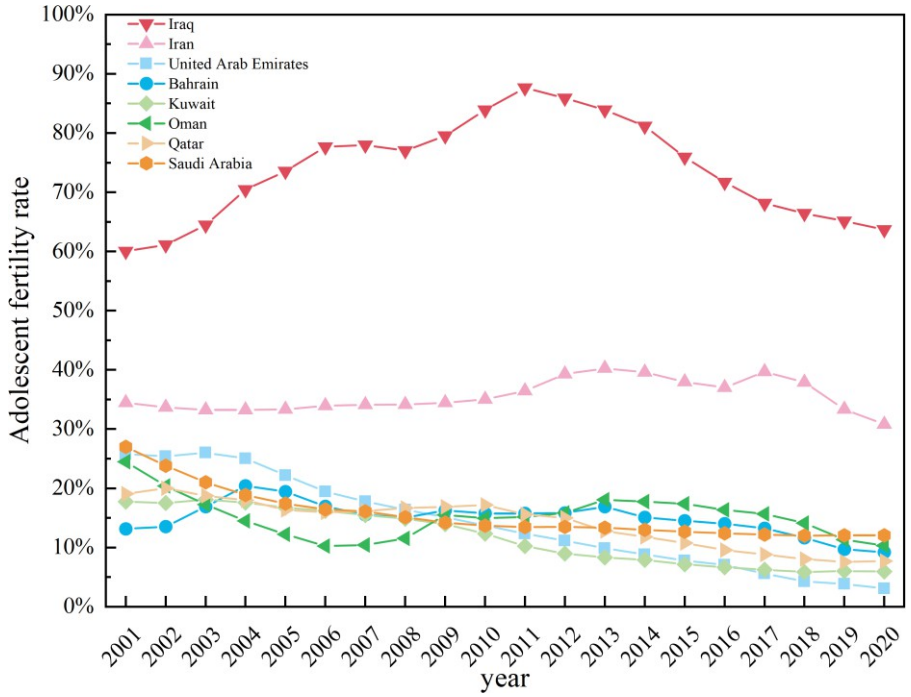
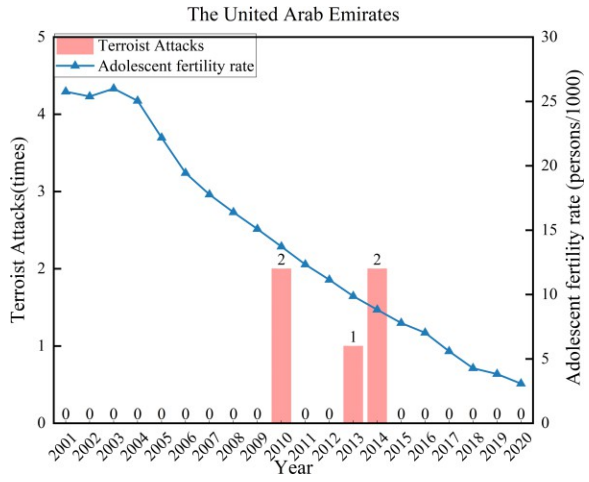


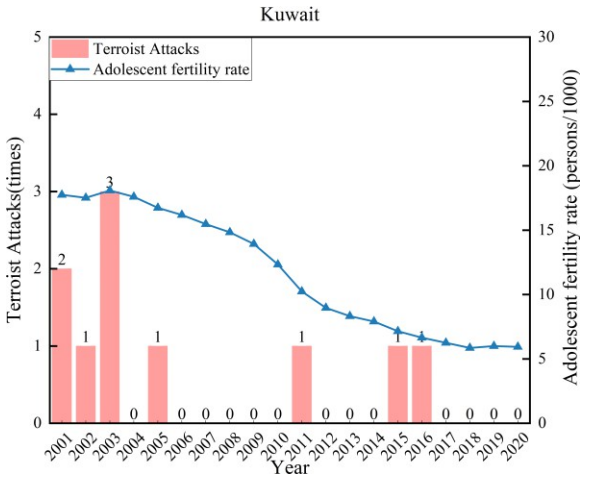
Figure 7. Comparison of adolescent fertility rates among the eight Gulf States

The downward trend in adolescent fertility rates over the years not only reflects the increasing emphasis on personal development and life choices but also demonstrates the deepening commitment to human rights protection. Figure 8(a)-(d) reveals that the United Arab Emirates, Kuwait, Oman and Qatar all experienced relatively few terrorist attacks between 2001 and 2020, with none exceeding 10 and an AFR of less than 30%. These countries also showed a steady downward trend. To some extent, this illustrates th e significant advancements these nations have made in upholding social stability and security amidst globalization, which assures individuals the opportunity to pursue higher-quality lives and personal growth. Within this relatively stable and secure social environment, individuals are able to make more autonomous and responsible choices in areas such as childbirth, education, and career, which is a tangible manifestation of human rights protection. Upon further examination of Figures 8(e)&(f), it becomes evident that the sustained occurrence or sudden increase in terrorist attacks, as observed in Iran and Bahrain, can potentially slow down or even reverse the downward trend in adolescent fertility rates. The situation in Saudi Arabia (Figure 8(g)) corroborates this observation.

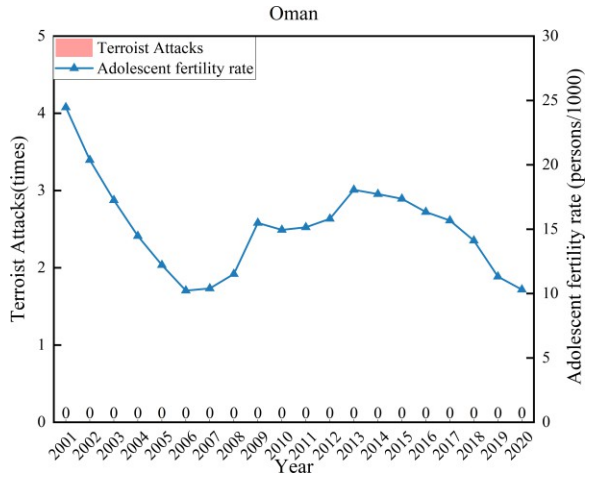
It is particularly noteworthy that the impact of terrorist attacks on AFR may manifest indirectly through various dimensions. Following a terrorist attack, members of society may be inclined to encourage earlier marriage and childbirth among adolescents due to a sense of insecurity. Economic recession may force adolescents into the labor market earlier, thus affecting their fertility plans. Restrictions on education and cultural activities may limit adolescents' educational and social opportunities, indirectly influencing their fertility decisions. Additionally, terrorist attacks may adversely affect the reproductive health of adolescents by compromising the provision of medical services. In the case of Iraq (Figure 8(h)), the average annual number of terrorist attacks in the country is as high as 1,360.6 in 2001-2020, and the AFR has remained above 60% over the 20 year. Specifically, between 2001 and 2012, as terrorist attacks escalated, the adolescent fertility rate also rose consistently. However, from 2017 to 2020, as the situation stabilized, the fertility rate began to decline. This phenomenon underscores the significant impact of heightened terrorist activity on social instability and public panic, ultimately driving up fertility rates. Conversely, when terrorist attacks are effectively contained or social stability is restored, fertility rates may respond by decreasing. This relationship highlights the crucial role of social stability and a secure environment in promoting responsible reproductive behavior, further emphasizing the close linkage between human rights protection and social development.



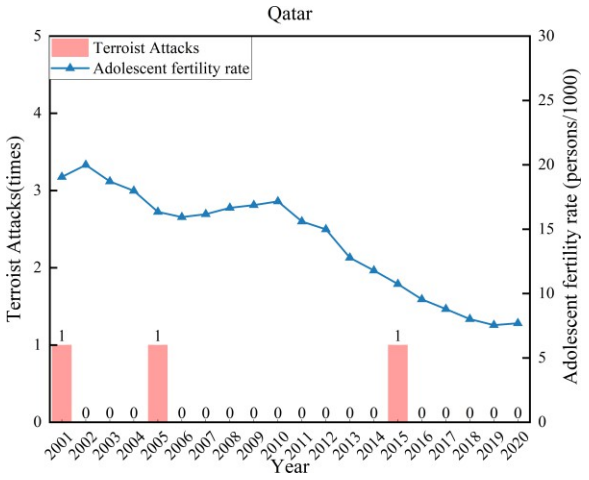
(a) The tendency of Terrorist Attacks and AFR in the United Arab Emirates



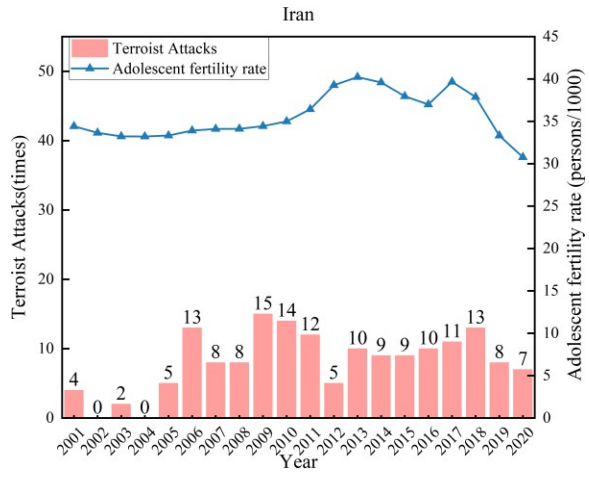
(b) The tendency of Terrorist Attacks and AFR in Kuwait



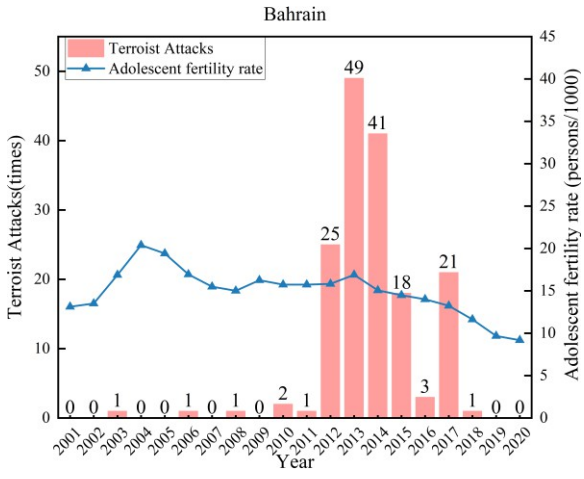
(c) The tendency of Terrorist Attacks and AFR in Oman



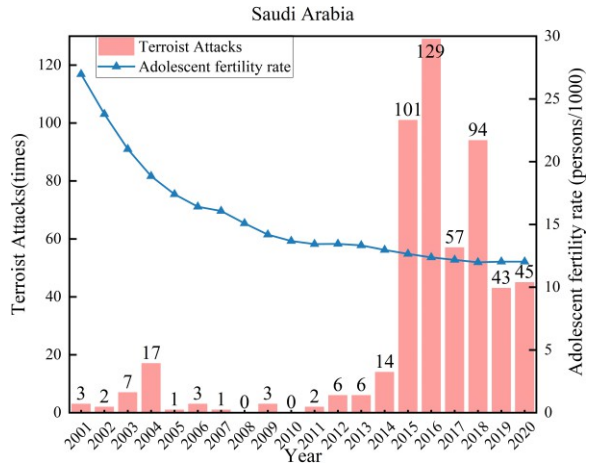
(d) The tendency of Terrorist Attacks and AFR in Qatar



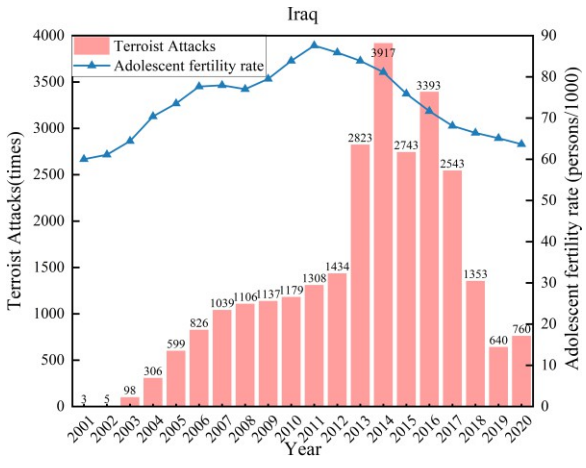
(e) The tendency of Terrorist Attacks and AFR in Iran



(f) The tendency of Terrorist Attacks and AFR in Bahrain



(g) The tendency of Terrorist Attacks and AFR in Saudi Arabia



(h) The tendency of Terrorist Attacks and AFR in Iraq

Figure 8. The tendency of Terrorist Attacks and AFR in Qatar in the eight Gulf States

LIMITATIONS AND FUTURE WORK

Firstly, this study aims to provide a correlation analysis between terrorist attacks and other socio-economic indicators, rather than causal analysis. The authors acknowledge that while causal analysis is more valuable, it is also more challenging to prove statistically whether a certain factor is the cause of terrorist attacks or a result of terrorist attacks. In addition, the mediators behind the changes in adolescent fertility, on which the article ultimately focuses, have not been thoroughly analyzed. In Figure 8, an attempt was made to illustrate the positive correlation between terrorist attacks and adolescent fertility rates through the analysis of individual countries, and qualitatively explored the underlying potential factors. However, this representation may have somewhat diminished the comparative effectiveness among countries. To enhance the depth and accuracy of the analysis, future research could delve deeper into two key directions: (1) whether differences in terrorist attack rates can serve as indicators of disparities in adolescent fertility rates among different countries, requiring comprehensive analysis and comparison of cross-national data; (2) how fluctuations in terrorist attack rates within a single country affect its domestic birth rate, which necessitates detailed national-level data support.

Secondly, as a case study, the authors only selected the Gulf states for analysis, consisting of eight countries. The authors believe that the threat of terrorism is becoming a serious international concern and requires a more comprehensive and larger-scale analysis. Moreover, the findings of this paper, which are based solely on the Gulf states, may not be applicable to other countries and regions globally. Therefore, the authors plan to expand the scope of the study by including more countries and regions to gain a better understanding of the global risk on terrorist attacks. In the near future, we may even conduct a global-scale analysis to uncover more universal and inspiring findings.

CONCLUSIONS

This paper presents a thorough analysis of terrorist attacks in Gulf states using the GTD database. The case study encompasses totally eight countries: Bahrain, Iran, Iraq, Kuwait, Qatar, Saudi Arabia, United Arab Emirates and Oman. As a preliminary exploratory study, Spearman's correlation analysis and Gray correlation analysis methods were employed to explore the correlation between terrorist attacks and potential socio-economic factors. The socio-economic factors considered in this paper include per capita GDP, unemployment rate, net migration, adolescent fertility rate, military strength, urban population, agricultural land, and commodity trade.

It was observed that:

(1) Between 2001 and 2020, the terrorism situation in the Gulf states experienced continuous fluctuation. From 2001 to 2010, the fluctuations in the terrorism situation in the Gulf states were relatively smooth. Although there were some major incidents, such as the wars in Iraq in 2003 and Lebanon in 2006, which heightened regional tensions, the overall trend was not significantly upward. From 2010 onwards, there was a fluctuating growth trend in terrorism in the Gulf states. Iraq bore the brunt of terrorist attacks. Since 2016, this trend has gradually weakened and the number of terrorist attacks decreased. However, terrorism remains a significant security risk that negatively affects economic development, overseas investments, and the well-being of the population.

(2) Terrorism in the Gulf states is a complex phenomenon resulting from multiple factors. This war-ravaged region has become a breeding ground for terrorism, further aggravating the fragile security situation in the Gulf states. The study concludes that the unemployment rate, arable land ratio, military strength and teenage fertility rate have significant positive correlations with terrorist attacks, while the urban population ratio, GDP per capita and commodity trade have significant negative correlations. These factors are interrelated and pose a significant threat to the security and stability of the Gulf and the Middle East region.

(3) Children residing in regions experiencing a high frequency of terrorist attacks often face immense difficulties in fully exercising their rights to education and. The prevalent insecurity and terrorist activities in these areas pose a profound danger to the survival and development of children and adolescents, effectively barring them from enjoying their fundamental rights to the fullest extent. Consequently, it is incumbent upon the international community and each individual nation to prioritize the survival and sustainable development of children and adolescents, implementing effective measures to safeguard their rights and welfare.

Future studies are expected to delve deeper into causal analyses of terrorist attacks and conduct comprehensive global-scale investigations. This paper proposes that the Gulf States expeditiously forge a consensus on regional security, vigorously facilitate security cooperation, and boost investments in safeguarding the welfare of children and adolescents. Additionally, the concerted effort of the entire international community is imperative to collectively tackle the humanitarian challenges posed by terrorism, attain sustainable regional and global security, and enhance the protection and development opportunities for all children and adolescents worldwide.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (Grant No. 72304038), Guangdong Basic and Applied Basic Research Foundation (Grant No. 2022A1515110049), Young Innovative Talents Project from Department of Education of Guangdong Province (Grant No. 2022KQNCX157).

REFERENCES

Braithwaite, A., & LI, Q. (2007). Transnational terrorism hot spots: Identification and impact evaluation. *Conflict Management and Peace Science*, 24(4), 281-296. https://doi.org/10.1080/07388940701643623

Buesa, M., Valiño, A., Heijs, J., Baumert, T. & Gómez, J.G. (2007). The economic cost of March 11: Measuring the direct economic cost of the terrorist attack on March 11, 2004 in Madrid. *Terrorism and Political Violence*, 19(4), 489-509, https://doi.org/10.1080/09546550701590677

Cho, S. (2010). A study on analysis and response against Middle East terrorism, since 1990s. *Journal of Korean Public Police and Security Studies*, 7(3), 89-110.

De Cauwer, H., Barten, D.G., Tin, D., Mortelmans, L.J., Ciottone, G.R., & Somville, F. (2022). Terrorist attacks against concerts and festivals: A review of 146 incidents in the global terrorism database. *Prehospital and Disaster Medicine*, 38(1), 1-8. https://doi.org/10.1017/S1049023X22002382

De Cauwer, H., Barten, D.G., Tin, D., Mortelmans, L.J., Ciottone, G.R., & Somville, F. (2023). 50 years of terrorism against the nuclear industry: A review of 91 incidents in the global terrorism database. *Prehospital and Disaster Medicine*, 38(2), 199-206. https://doi.org/10.1017/S1049023X2300002X

Deng, S.C., Qin, Y., Xu, J.F., & Gu, Y.W. (2021). Quantitative classification of global terrorist attacks and prediction of anti-terrorist situation: Based on projection pursuit model and grey metabolic model. *Science & Technology Review*, 39(18), 111-121. [in Chinese]

Global Terrorism Database (GTD). (2023). https://www.start.umd.edu/gtd/

Holman, E.A., Silver, R.C., Poulin, M., Andersen, J., Gil-Rivas, V., & McIntosh, D.N. (2008). Terrorism, acute stress, and cardiovascular health: A 3-year national study following the September 11th attacks. *Arch Gen Psychiatry*, 65(1), 73–80. https://doi.org/10.1001/archgenpsychiatry.2007.6

Kim, W., & Sandler, T. (2020). Middle East and North Africa: Terrorism and conflicts. Global Policy, 11(4), 424-438. https://doi.org/10.1111/1758-5899.12829

Krueger, A.B., & Malečková, J. (2009). Attitudes and action: Public opinion and the occurrence of international terrorism, Science, 325(5947), 1534-1536. https://doi.org/10.1126/science.1170867

Kydd, A. H., & Walter, B. F. (2006). The strategies of terrorism. *International security*, *31*(1), 49-80.

Leite, A., Ramires, A., Dinis, M.A.P., Sousa, H.F.P.E. (2019). Who is concerned about terrorist attacks? A religious profile. *Social Science-Basel*, 8(11), 316. https://doi.org/10.3390/socsci8110316

Li, Z., Li, X., Dong, C., Guo, F., Zhang, F., & Zhang, Q. (2021) Quantitative analysis of global terrorist attacks based on the global terrorism database. *Sustainability*, 13, 7598. https://doi.org/10.3390/su13147598

Morris, N.A., Lafree, G., Karlidag, E. (2021). Counter-terrorism policies in the Middle East: Why democracy has failed to reduce terrorism in the Middle East and why protecting human rights might be more successful. Criminology & Public Policy, 20(1), 153-175. https://doi.org/10.1111/1745-9133.12532

Mun, S., Han, S.H., & Kim, H.J. (2021). Terrorist attacks and total factor productivity. *Economics Letters*, 202, 109808. https://doi.org/10.1016/j.econlet.2021.109808

Pan, X.H. (2021). Quantitative analysis and prediction of global terrorist attacks based on machine learning. *Scientific Programming*, 2021, 7890923. https://doi.org/10.1155/2021/7890923

Piazza, J.A. (2007). Draining the swamp: Democracy promotion, state failure, and terrorism in 19 Middle Eastern countries. *Studies in Conflict & Terrorism*, 30(6), 521-539. https://doi.org/10.1080/10576100701329576

Richards, A. (2014). Conceptualizing terrorism. *Studies in Conflict & Terrorism*, *37*(3), 213-236.

Roy, N., Kapil, V., Subbarao, I., & Ashkenazi, I. (2011). Mass casualty response in the 2008 Mumbai terrorist attacks. *Disaster Medicine and Public Health Preparedness*, 5(4), 273-279. https://doi.org/10.1001/dmp.2011.80

Sharma, D. (2013). Growing overlap between terrorism and organized crime in India: A case study. *Security Journal*, 26, 60-79. https://doi.org/10.1057/sj.2011.33

Tin, D., Fares, S., Al Mulhim, M., & Ciottone, G.R. (2022). Terrorist attacks in the Middle East: A counter-terrorism medicine analysis. *Prehospital and Disaster Medicine*, 37(2), 212-216. https://doi.org/10.1017/S1049023X22000358

Tong, Y.-S., Pang, Y., & Yang, F.-D. (2016). Psychological first aids after September 11 terrorist attack. *Chinese Mental Health Journal*, 30(10), 775–778.

Yaya, M. E. (2009). Terrorism and tourism: The cause of Turkey. *Defence and Peace Economics*, 20(6), 477-497. https://doi.org/10.1080/10242690903105414

Ulmer, N., Barten, D.G., De Cauwer, H., Gaakeer, M.I., Klokman, V.W., van der Lugt, M., Mortelmans, L.J., van Osch, F.H.M., Tan, E.C.T.H., Boin, A. (2022). Terrorist attacks against hospitals: World-wide trends and attack types. *Prehospital and Disaster Medicine*, 37(1), 25-32. https://doi.org/10.1017/S1049023X22000012

Varenia, N., Ryzhov, I., Avdoshyn, I., Hrechin, Y., Rozvadovskyi, O. (2021). The phenomena of radicalism and extremism: Socio-legal hypotheses for the response, AD ALTA-Journal of Interdisciplinary Research, 11(2), 199-203.

Understanding Crisis Preparedness: Insights from Personal Values, Beliefs, Social and Personal Norms

Felix Burger
Technical University of Darmstadt, FAI
felix.burger@magenta.de

Jasmin Haunschild
Technical University of Darmstadt, PEASEC
haunschild@peasec.tu-darmstadt.de

Christian Reuter
Technical University of Darmstadt, PEASEC
reuter@peasec.tu-darmstadt.de

ABSTRACT

In disasters and crises, crisis preparation at the household level is crucial for societal resilience. However, motivating citizens to take personal responsibility and to implement preparedness measures remains challenging. This study investigates the impact of values, beliefs and norms on preparedness. Through a survey in Germany on resource-oriented crisis preparedness and participants' values, a group of more prepared and less prepared individuals was differentiated. We quantitatively analyzed the correlation between values and crisis preparedness. A quantitative analysis revealed a significant association between the value of *benevolence* and higher preparedness. Follow-up interviews (N=30) explored participants' beliefs and norms, showing minimal group differences. Although crisis preparedness is generally viewed positively, strong negativity exists towards perceived "hoarders". The study discusses implication for crisis communication and the design of nudges and persuasive systems, emphasizing intervention that can trigger benevolence to foster preparedness and appeal to the social benefits to counter the narrative of anti-social crisis preparedness.

Keywords

Crisis preparedness, values, beliefs, social norms, personal norms

INTRODUCTION

Despite advancements in understanding the causes and consequences of disasters, they continue to result in significant harms and losses. Climate change is leading to more frequent and severe natural disasters. Recently, crises such as the global Covid-19 pandemic and the invasion of Russia in Ukraine have demonstrated the importance of societal resilience and crisis preparedness. Crisis preparedness is defined as the knowledge and skills enabling anticipation, response, and recovery from disasters (UNDRR, 2016). This applies across various levels, from individual households to international organizations.

Individual households play a crucial role as even well-prepared regions may wait up to 72 hours for external assistance in disaster situations (Kohn et al., 2012). In addition, due to the increasing dependence of infrastructure, energy and internet outages can have severe negative impacts on civil society (UP KRITIS, 2014). Therefore, civil protection and disaster relief agencies often promote household crisis preparedness. They commonly recommend water and food supplies, energy sources, information access, hygiene maintenance, medical care, document copies, fire protection, and emergency kits (BBK, 2019). However, in many countries, citizens perceive that they are badly informed and unprepared for crises (Appleby-Arnold & Brockdorff, 2018) and many countries note insufficient crisis preparedness (Garschagen et al., 2019; Menski, 2016). Research suggests that families and retirees are better prepared, whereas single households, students/apprentices and people with low financial power

lag behind (Gerhold et al., 2019). Convenience appears to be more significant for preparation than risk awareness (BBK, 2013; Gerhold et al., 2019). In addition, urban residents, possibly due to the lack of space, are less prepared and more dependent on infrastructure (BBK, 2013).

Psychological factors have also been found to influence crisis preparedness behavior, particularly by affecting risk perception. Unrealistic optimism and risk compensation are widespread. In some risk cultures, there is often an overreliance on state agencies to handle crises, affecting individual preparation (Cornia et al., 2016). This indicates that prevailing beliefs about how crises occur and are managed, and prevailing social norms might play a role.

In addition, values influence behavior through norms and beliefs (Schwartz, 1992; Steg et al., 2005). As such, psychological motivation theories, such as the Values-Beliefs-Norms theory (VBN-theory), are often based on the users' values. Values are "desirable transsituational goals varying in importance, which serve as a guiding principle [...]" (Schwartz, 1992, p. 21). They are thus more overarching and stable than beliefs and have been shown to impact behavior related to sustainability, altruism and health (Honka et al., 2019; Schwartz, 2016; Steg et al., 2005). In decision-making and goal-setting values help people define and consider what is meaningful and worthwhile and by shaping moral judgments and actions (Schwartz, 1992). Values impact behavior directly and indirectly through personal norms and beliefs (Steg et al., 2005; Stern, 2000).

Finally, social norms influence behavior (Reynolds et al., 2015). Social norms are unwritten rules for specific behavior that are maintained through group approval or punishment. They manifest, for example, in an individual's perception that most people important to them should or should not do something (Ajzen & Fishbein, 1980; Kim & Seock, 2019). These norms, alongside values, have been shown to influence personal norms and, consequently, have both indirect and direct effects on concrete behavior (Kim & Seock, 2019).

However, the roles of values, beliefs, personal and social norms have so far not been explored with regard to crisis preparedness. This study contributes a first exploration by analyzing differences in these four aspects between people in Germany (N=30) who are more and less prepared for crises. After presenting the state of research concerning crisis preparedness and the impact of values, beliefs, personal and social norms on behavior, we describe the methodologies used. Then, we present the findings from a quantitative survey investigating preparedness and values, and qualitative interviews that explored beliefs and norms. Finally, we discuss the findings, before summarizing the study in the conclusion.

RELATED WORK

Crisis Preparedness

Despite the primary responsibility of state and volunteer agencies for emergency management, individual and household crisis preparedness remains crucial, especially in case of large-scale crises or infrastructure failures. Since agencies cannot serve everybody immediately, states give out recommendations for individual and household crisis preparedness. However, factors influencing the implementation of such recommendation and household crisis preparedness are currently not well-understood. Limited data on disaster preparedness exists and little research has explored the individual preparedness of vulnerable population groups (Garschagen & Sandholz, 2018), thus precise reasons for lower crisis preparedness remain unknown. Existing literature suggests factors such as household size, presence of children, square meters per person in the dwelling, educational level, employment situation, and urbanization level in inhabitants per square kilometer as relevant to preparedness levels. Furthermore, nationality may play a role due to lower experience regarding local crises, reduced awareness of informational resources, and weaker social networks potentially contributing to lower preparedness.

While families and retired individuals seem well-prepared in terms of resources, single households and those of students or apprentices often lack the recommended provisions, making them dependent on external assistance (Gerhold et al., 2019). Interestingly, despite better preparation, risk perception decreases with age (Gerhold et al., 2019), suggesting that financial resources and available living space might be important limiting factors. Other studies conclude that household size, educational level or prior crisis experiences have little impact on individual preparedness, instead, well-prepared groups might be better equipped for the sake of convenience (Garschagen et al., 2019). Urbanization also plays a role, with individuals in cities being less prepared and, consequently, more reliant on the extensive infrastructure (BBK, 2013). Particularly in urban areas, extended critical infrastructure failures, such as shortages in drinking water, can lead to severe conditions, given that only half of the surveyed population has water reserves (Garschagen et al., 2019).

Psychological Constructs and Crisis Preparedness

Alongside demographic factors and living condition, psychological factors influence crisis preparedness and risk perception. The VBN theory posits a causal chain among values, behavior-specific beliefs, personal norms, and behavior (Stern, 2000). While initially designed for environmentally conscious actions, it is believed to also be applicable to other areas of behavioral research (Steg et al., 2005).

Values, having the highest explanatory power, initiate the chain. Behavior is influenced by personal norms, which refer to individual ethical guidelines or self-imposed expectations and obligations. Affected by the general opinion within a group, personal norms can differ from it, particularly influencing altruistic behaviors and others as internalized values (Schwartz, 1977; Thøgersen & Ölander, 2002). Personal norms, in turn, are influenced by behavior-specific beliefs, fundamental thoughts about a specific behavior, affecting how it's perceived in terms of difficulty and facilitators or hindrances e.g., the conviction that there are situations threatening something important to the individual and the belief that the person can reduce this threat through actions. These beliefs, in turn, can be explained by certain values, such as "security" or "benevolence" (Steg et al., 2005; Stern, 2000). In the context of environmental protection, values reflecting concern for non-human species or the biosphere (e.g., "universalism") play a significant role. Conversely, self-centered and conservative values (e.g., "power") are negatively correlated with environmental friendliness (Steg et al., 2005). It is plausible that the value of "security" is related to individual crisis preparedness behavior, although the literature does not specify this connection.

Values "(1) are concepts or beliefs, (2) pertain to desirable end states or behaviors, (3) transcend specific situations, (4) guide selection or evaluation of behavior and events, and (5) are ordered by relative importance" (Schwartz, 1992, p. 4). Twelve internationally universal values have been identified that can be clustered into four categories (Grosz et al., 2021; building on Schwartz et al., 2012). These values serve as guiding principles for individuals, influencing their attitudes, actions, and decision-making. Individuals prioritize values to varying degrees, creating their unique value profiles. They also shape people's personal norms, attitudes, preferences, and worldview. While values are broad, enduring beliefs that guide overall behavior, personal norms are more specific, context-dependent guidelines that help shape behavior in particular situations. As such, values are antecedents to personal norms and help explain them (Steg et al., 2011). Some values have been identified to be more often prioritized together, while others are typically more opposed (see Figure 1).

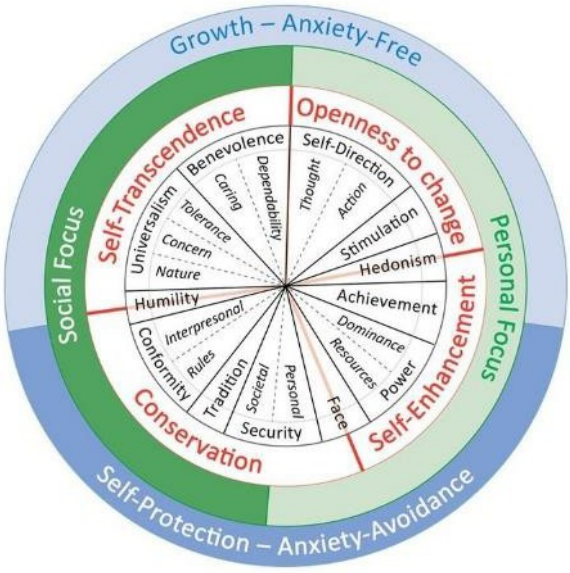


Figure 1. Value circumplex showing 12 basic values, 19 personal values, four higher-order values, and two underlying motivations. Source: Grosz et al., 2021; building on Schwartz et al., 2012. Licensed under CC-BY 4.0.

The opposition of *self-transcendence* versus *self-enhancement* reflects the tension between valuing the welfare of others and the broader society versus pursuing one's own interests and success. Self-transcendence includes values such as universalism (promoting understanding, tolerance, and protection of the welfare of all people) and benevolence (caring for others). By contrast, self-enhancement contains values such as power (seeking control over others), achievement (striving for personal success), and hedonism (pursuit of pleasure and enjoyment). The second opposition, openness to change versus conservation reflects the tension between valuing novelty, change, and exploration versus valuing tradition, stability, and preservation of the existing order. Openness to change

includes values such as self-direction (choosing one's own goals and actions), stimulation (seeking excitement and new experiences), and hedonism (pursuit of pleasure and enjoyment). Opposed to this, conservation contains values such as security (seeking safety and stability), tradition (respecting customs and traditions), and conformity (complying with social norms and expectations).

Past research has shown that environmental factors, such as time constraints, decrease the power of values, reducing pro-value behavior, e.g., the likelihood of offering help (Latané & Darley, 1976). Additionally, the impact of values on behavior can be heightened by activating values, and by a behavior being perceived as relevant to the value (Schwartz, 2016).

Cognitive distortions, such as unrealistic optimism, often lead individuals to underestimate their personal vulnerability to crises compared to peers, resulting in insufficient preparation (Weinstein, 1980). Additionally, people tend to engage in risk compensation, adjusting their preparedness behavior based on the subjectively perceived level of threat in their environment (Etkin, 1999). Therefore, visible government-led crisis management measures can reduce individuals' perceived threat, leading to decreased individual preparedness. In some risk cultures, particularly “state-oriented” and “fatalistic” risk-cultures, the populations tend to believe that crises cannot be averted, or that the responsibility for management lies with the government. This diminishes citizens’ motivation for personal preparation, leading to relatively poor preparedness and low confidence in the ability to independently cope with crisis situations (Cornia et al., 2016).

Research Gap

Values, beliefs and norms and related behavior change interventions have so far primarily been investigated with regard to environmentalism and health (Canlas et al., 2022, Nwafor et al., 2021). Only two studies looked at nudging and preparedness: One focused on a single intervention for household flood preparedness (Mol et al., 2021), finding no advantage of their social norm-nudges regarding readiness to prepare. Another has empirically tested the effect of a social and a rational-confrontational nudge on crisis preparedness, finding that only the latter, confronting with the costs and benefits of preparedness, increased the taking of crisis preparedness measures (Haunschild et al., 2023). While some studies have demonstrated that culture shapes crisis preparedness (Appleby-Arnold et al., 2021; Cornia et al., 2016), they have not engaged with underlying values, beliefs or norms.

Still, currently, no research exists on the role specifically of beliefs, norms and values for crisis preparedness. Therefore, we question the impact of each in the following research questions. First, we examine the role of different *values* by asking the question:

- **RQ1:** Do well-prepared individuals differ from less-prepared ones regarding their *value orientations*?

Personal norms play a significant role as predictors of behavior within the VBN theory (Stern, 2000). A study investigating relevant factors for flood preparedness suggests that personal norms may be significant (Mol et al., 2021), leading to RQ2:

- **RQ2:** Do well-prepared individuals differ from less-prepared ones with regard to their *personal norms* regarding crisis preparedness?

A previous study analyzing the use of warning apps has found that expected efficacy of the response significantly influences the adoption and continued use of warning apps (Fischer-Preßler et al., 2022). In addition, in most countries, disaster worry is related to the intention to take preparedness measures, whereas in Germany, perceived disaster risk is more relevant (Appleby-Arnold & Brockdorff, 2018). Another branch of research indicates that countries differ with regard to their risk culture and the resulting perceived personal responsibility (Cornia et al., 2016). Therefore, *beliefs* about the effectiveness of individual preparedness measures may influence individual crisis preparedness intentions.

- **RQ3:** Do well-prepared individuals differ from less-prepared ones with regard to their *beliefs* about crisis preparedness?

Perceived *social norms* are generally considered a strong motivator for behavior (Goldstein & Cialdini, 2007). This has also been shown to be relevant for measures related to crisis response, such as the adoption and continued use of warning apps (Fischer-Preßler et al., 2022). However, a study investigating household flood preparedness did not find this to be relevant (Mol et al., 2021), possibly due to the relative invisibility of anti-flood investments or the relatively high costs. Therefore, we explore the unclear role of social norms for less costly household preparedness measures by answering RQ4:

- **RQ4:** Do well-prepared individuals differ from less-prepared ones regarding their *perceived social norms* surrounding crisis preparedness?

METHOD

Recruitment and Sample

To recruit the sample, the study was advertised via social media platforms, including Facebook groups focusing on weather, police, and traffic updates, student *WhatsApp* instant messenger groups, the survey website *SurveyCircle* and a university email list. To ensure a diverse sample, encompassing individuals with varying levels of preparedness across demographic characteristics, participants from underrepresented groups were asked to motivate additional participants. Table 1 describes the sample. It shows that gender (with 47% females in the sample compared with 50.7% in Germany (Statista, 2023b)) and the degree of urbanization are well-represented (80% of participants living in cities compared to 78% in Germany (Worldbank, 2023). In terms of citizenship, with 23% foreigners the sample is similar to the German population (with 16% foreigners (Federal Statistical Office, 2023a). Household net income categories are based on the inflation-adjusted categorization from Gerhold et al. (2019) and are well-represented in the sample (Bundeszentrale für politische Bildung, 2020). The age median is in the group of 26-30-year-olds, presenting a section younger than the general population (*MD* = 45 years (Statista, 2023a)). The sample over-represents academics, with 60% having an academic degree compared to 18,5% in the German population (Federal Statistical Office, 2020). The distribution of household size shows fewer single- and more three-person households in the sample compared to the population (Federal Statistical Office, 2023b).

Table 1. Sample Description

Gender	Female: 13, Male: 17
Age	21-25 years: 10, 26-30 years: 8, 31-35 years: 3, 36-40 years: 1, 41-50 years: 3, 51-60 years: 2, 61-70 years: 3
Formal Educational	No degree: 2, Secondary or tertiary education: 6, Vocational training: 4, Bachelor's degree: 7, Master's degree or higher: 11
Household Size	Single: 6, Two persons: 8, Three persons: 11 (2 with children), Four persons: 2, Five persons: 1, Six persons: 1, Seven persons: 1
Household Net Income	< 1699€: 10, 1700-3000€: 7, > 3000€: 13
Urbanization	Rural (<1000 persons/km²): 6, Rural-urban (1001-2000 persons/km²): 15, Urban (>2000 persons/km²): 9
Citizenship	German: 23, European: 3, Non-European: 4

Data Collection and Analysis

Data was captured via an online questionnaire and semi-structured follow-up interviews. The survey was conducted online using the GDPR-compliant software *SoSci-Survey*. The average completion time was 20 minutes. The first of three parts gathered individual socio-demographic data (i.e., gender, age, nationality, educational qualification, employment) and data on the participant's household (i.e., size, living space, number of children, net income). For the second part, a definition of crisis preparedness was presented, and participants responded to twelve questions about their preparedness levels. These questions were based on recommendations by the Federal Office of Civil Protection and Disaster Assistance and included primarily resource-oriented preparation, that would be relatively cheap and easy to obtain, such as food, water, first aid and hygiene supplies, fire protection, energy and information sources (batteries, radio) and document backups (BBK, 2019). The third part of the questionnaire assessed values associated with crisis preparedness behavior by using the Portrait Values Questionnaire (PVQ) (Schwartz et al., 2012), adapted to crisis preparedness. For example, for the value "power", participants stated their agreement with the statement "He engages in crisis preparedness because the sense of control that resources can provide is important to him" (adjusted to participants' gender). The adapted PVQ demonstrated acceptable to excellent reliability based on assessments of internal consistency (Cronbach's α between .70 and .96).

A crisis preparedness score was computed by assigning 0 to 3 points to participants based on their responses to various preparedness measures. The assigned value of the measures was based on the risk index by (Nohl & Thiemecke, 1988) which takes the potential negative cost of the missing measure into account (see Table 2). The score was normalized to a measure between 0 and 1 by dividing the score by its maximum of achievable points.

Table 2. Assignment of Preparedness Score Values

Preparedness measures	**Points**			
	0	1	2	3
Number of days household ... has sufficient food supply is sufficient	0+	3+	7+	10+
... water supply	0+	3+	7+	10+
... has access to light/heat/cooking (energy sources)	0+	3+	7+	10+
... can access news	0+	3+	7+	10+
...can maintain hygiene standards	0+	3+	7+	10+
Time since the last first aid course	> 5 years	2-5 years	< 2 years	< 6 months
Presence of a first aid kit	No	Yes	-	-
Presence of medications	No	Yes	-	-
Knowledge of the location of important documents & readily available/digital copies	No	Knowledge	Copies	Knowledge + copies
Fire protection (fire blanket or fire extinguisher)	No	-	Yes	-
Emergency supplies (bug-out bag)	No	-	Yes	-

The median (*MD*=0.45, *SD*=0.16, spanning from .11 to .78) was used as the baseline to differentiate the groups of rather more (Group 1, N=16) and rather less prepared individuals (Group 2, N=14). Following standard practice (Schwartz & Cieciuch, 2022), the scores for the individual values were computed by averaging the related item scores. To explore the relationship between crisis preparedness behavior and personal values, Mann-Whitney-U-tests for comparing independent groups with non-normally distributed dependent variables, were performed using R statistical software. The interview (see Table 3) started with creating a comfortable environment to encourage open responses, then explored participants' understanding of crisis and preparedness, and offered a definition of crisis preparedness (UNDRR, 2016). Next, participants' beliefs, social norms and personal norms related to the topic were discussed. Interviews were conducted remotely between June 26 and July 19, 2023 and were recorded with participant consent, lasting an average of 25 minutes. The questionnaire (N=5) and interview (N=1) were pretested, leading to minor adjustments.

Table 3. Interview Guide

(1) Introduction and establishment of trustful atmosphere

(2) Topic introduction
- Do you engage in crisis preparation?
- What does the term "crisis preparation" mean to you?
- How did you come to (not) prepare yourself?

(3) Beliefs
- What advantages do you see in preparing for crises? What about disadvantages?
- What makes it difficult/easy to prepare for crisis situations?
- What motivated you to buy X/prepare for Y, or what would motivate you to buy X/prepare for Y?

(4) Social norm
- Is crisis preparedness discussed in your environment? If yes, what does it look like? Are practical strategies for dealing with it discussed?
- Do you think other people in your environment prepare for crisis situations? If yes, how?
- Who do you think would approve if you prepare for crisis situations? Who would disapprove?

(5) Personal norm
- What do you think is the right approach to crisis preparedness for you?
- Regardless of what people in your environment think or do, how do you think you should behave?
- What are the characteristics of a person who prepares for crisis situations?

(6) Closing questions
- Do you have any additional points you would like to mention? Do you have any questions?

The interviews were pseudomized, transcribed and the records deleted. The interview data were analyzed using qualitative content analysis (Mayring, 2015), inductively building a coding scheme with relevant themes that emerged from the text analysis (see Figure 2).

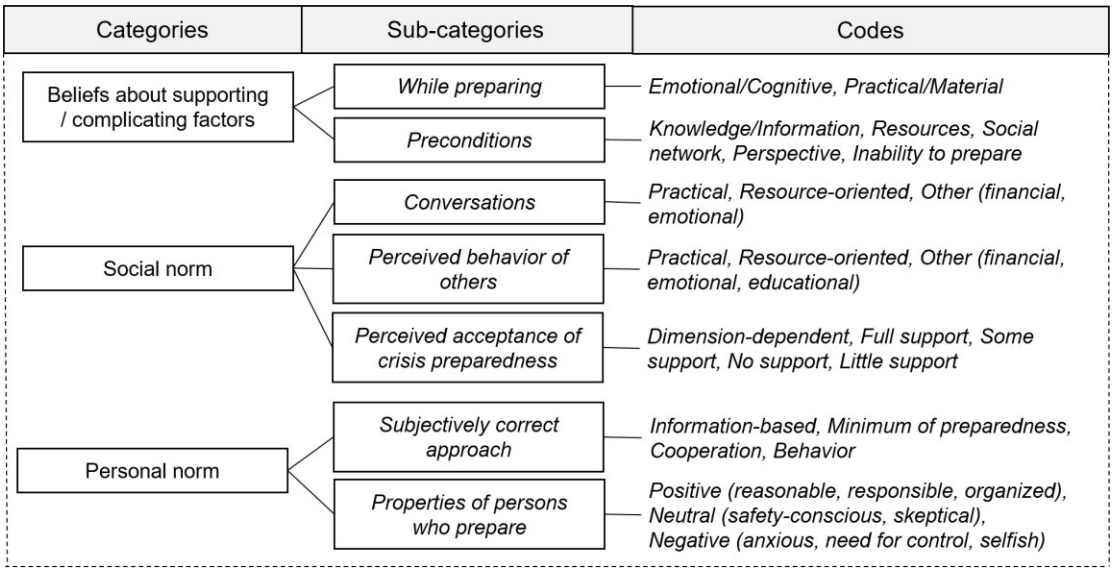


Figure 2. Coding System

RESULTS

Crisis Preparedness and Values (RQ1): Relevance of Value *Benevolence*

The prerequisites for using t-tests to address the research question were not met due to the non-metric scale level of the crisis preparedness score. Moreover, some value expressions did not fulfill the requirement of normal distribution and variance equality. Therefore, Mann-Whitney-U-tests were employed to examine possible differences in values between the differently prepared groups. Table 4 presents the means, standard deviations, and medians of values in the groups regarding crisis preparedness behavior, as well as the Mann-Whitney-U-statistic, Z-statistic, and p-values. A significant difference exists regarding the value of *benevolence*, which is higher among the individuals in the more-prepared group. An effect size of r=-.39 (Fritz et al., 2012) can be assumed, indicating a moderate correlation (Cohen, 1988).

Table 4. Comparisons between Differently Prepared Groups' Attitudes Regarding Crisis Preparedness Behavior, *p < .05

	Well-prepared (G1)			*Less-prepared (G2)*			*Mann-Whitney-U*	*Z-statistic*	*p-value*
	M	*SD*	*MD*	*M*	*SD*	*MD*			
Self-Direction	3.29	0.89	3.40	2.65	1.36	2.75	78.5	-1.39261	.17
Stimulation	2.67	1.22	2.85	2.02	1.11	1.70	80	-1.33026	.18
Hedonism	2.40	1.03	2.15	2.00	1.06	1.85	85.5	-1.10162	.27
Achievement	2.22	1.13	2.30	1.56	0.63	1.50	74	-1.57968	.10
Power	2.15	0.85	2.20	1.55	0.62	1.30	66	-1.91225	.06
Face	2.46	1.07	2.30	1.79	0.80	1.70	70	-1.74596	.08
Security	3.48	0.69	3.50	3.11	1.34	3.00	90.5	-0.89377	.38
Tradition	2.21	1.05	2.00	1.65	1.20	1.00	71	-1.70439	.07
Conformity	2.07	0.96	1.80	1.49	0.48	1.40	69.5	-1.76675	.08
Humility	2.62	1.07	2.50	2.40	1.16	2.30	97.5	-0.60277	.56
Benevolence	3.76	0.91	4.00	2.74	1.38	3.15	60.5	-2.14088	.03*
Universalism	3.31	1.15	3.25	2.62	1.47	2.50	78.5	-1.39261	.17

Crisis Preparedness and Personal Norms (RQ2): Well-Prepared Group Puts More Emphasis on Minimal Preparation

The personal norm was assessed through questions about the subjectively correct approach to crisis preparedness and the characteristics of individuals engaging in crisis preparedness (see Table 5 for an overview of the results**)**.

Perceived personal norm by more-prepared individuals: Concerning the subjectively correct approach, 75% (n=12) stressed information-based measures, with an emphasis on staying informed (n=7) and preparing plans (n=4). Many made statements about the extent of preparation (81%, n=13), primarily focusing on covering the essentials (n=11). People who prepare were described positively as being reasonable, forward-thinking, calm, intelligent, and communicative (75%). Neutral characteristics covered safety-consciousness and skepticism (44%), while negative traits included anxiety, a desire for control, social weakness, and occasional perceptions of being annoying (63%).

Perceived personal norm by less-prepared individuals: An information-based strategy was also emphasized by 79% of Group 2 (n=11), involving creating plans (n=5), staying informed (n=4), and following existing recommendations (n=2). Fewer participants discussed the extent of preparation (50%, n=7), suggesting a minimalistic approach (n=5). Positive attributes mentioned by Group 2 included being organized, responsible, and reasonable (71%). Neutral characteristics comprised safety-consciousness and admiration-seeking (36%), while negative traits comprised anxiety, pessimism, egoism, and mistrust (71%).

Table 5. Personal Norms , Group 1 = more prepared (n=16), Group 2 = less prepared (n=14)

	Group 1		Group 2	
	n	**%**	**n**	**%**
Subjectively correct approach to crisis preparedness				
Information based	12	75%	11	79%
Being informed	*7*	*44%*	*5*	*36%*
Making plans	*4*	*25%*	*4*	*29%*
Following recommendations	*1*	*6%*	*2*	*14%*
Minimal preparation	11	69%	5	36%
Cooperation based	1	6%	4	29%
Behavior based	4	25%	3	21%
Staying healthy	*2*	*13%*	*2*	*14%*
Do nothing	*1*	*6%*	*2*	*14%*
Properties of persons who prepare				
Positive	12	75% (reasonable, forward-thinking, calm, intelligent, communicative)	10	71% (organized, responsible, reasonable)
Neutral	7	44% (safety-conscious, skeptical)	5	36% (safety-conscious, admiration-seeking)
Negative	10	63% (anxious, desiring control, socially weak, annoying)	10	71% (anxious, pessimistic, egoistic, mistrustful)

Crisis Preparedness and Beliefs (RQ3): Less Knowledge and Fewer Perceived Benefits among the Less-Prepared

Participants' beliefs were assessed though their beliefs about disadvantages (see Table 6) and advantages (see Table 7) of crisis preparedness behavior, and perceived obstacles and facilitating factors.

Beliefs about benefits and obstacles by well-prepared individuals: 63% emphasized emotional-cognitive benefits during preparation, highlighting a sense of security (n=5), competence (n=5), and an opportunity for hypothetical thinking (n=2). A majority (56%, n=9) mentioned discomfort related to engaging with the topic, such as mental strain (n=4) and fear (n=3). Half of the group (n=8, 50%) indicated practical-material benefits, including better crisis management (n=7), such as being able to react faster. Practical-material drawbacks comprised the perceived effort (n=9), wastefulness (n=5), lack of guaranteed returns (n=3), and concerns about fostering competition (n=2). 13% (n=2) attributed no benefits to crisis preparedness behavior.

As preconditions, 69% (n=11) highlighted the importance of knowledge and information, with visibility (n=3), quantity (n=3), and recommendations (n=2) playing a crucial role besides having a plan (n=2). Additionally, 44% (n=7) pointed to needing a resource surplus, including money, space, time, and physical health. Social networks (31%, n=5) and a positive perspective on crisis preparedness (19%, n=3) were also mentioned. Hindering factors

comprised lack of knowledge or information (25%, n=4), specifically a lack of orientation, resource deficiency (50%, n=8), and a negative perspective on crisis preparation (81%, n=13) expressing the demotivating aspect of the unpredictability of crises (n=6), the lack of fun and perceived effort (n=5) and not rating preparedness valuable enough (n=2). Remarkably, 63% (n=10) believed that nothing would impede crisis preparedness behavior. However, 50% mentioned barriers dependent on the extent of preparation, such as high costs (n=8) and the lack of infrastructure for certain measures (n=2).

Table 6. Beliefs about Disadvantages, Group 1 = more prepared (n=16), Group 2 = less prepared (n=14)

	Group 1		Group 2	
	N	**%**	**N**	**%**
While preparing				
Emotional-cognitive	9	56%	8	57%
Worry/fears	*3*	*19%*	*3*	*21%*
Mental strain	*4*	*25%*	-	
Stress/anxiety	-		*3*	*21%*
Embarrassment	-		*1*	*7%*
Unpredictability	*1*	*6%*	-	
Practical-material	15	94%	13	93%
Investments	*9*	*56%*	*10*	*71%*
Money	*2*	*13%*	*5*	*36%*
Time	*2*	*13%*	*3*	*21%*
Space	*2*	*13%*	*3*	*21%*
Repeated effort	*2*	*13%*	*1*	*7%*
Physical	-		*1*	*7%*
Waste	*5*	*31%*	*6*	*42%*
Competition	*2*	*13%*	*3*	*21%*
Uncertain gains	*3*	*19%*	-	
None	-		1	7%
Preconditions				
Lack of knowledge/information	4	25%	8	57%
Lack of orientation	*4*	*25%*	*4*	*29%*
Lack of topic visibility	-		*3*	*21%*
Lack of resources	8	50%	5	36%
Space	*5*	*36%*	*3*	*21%*
Money	*3*	*19%*	*2*	*14%*
Time	*3*	*19%*	*2*	*14%*
Physical	*2*	*13%*	*1*	*7%*
Attitude	13	81%	7	50%
Low gain expectancy	*2*	*13%*	*4*	*29%*
Unpredictability	*6*	*38%*	*2*	*14%*
Not feeling like it/exhausting	*5*	*31%*	*1*	*7%*
None	-		1	7%
What makes it impossible?				
Nothing	10	63%	8	57%
Limitations (disability, illness)	1	6%	4	29%
Lack of awareness	1	6%	-	
Perfect preparedness impossible	8	50%	3	21%
Insufficient supply infrastructure	2	13%	-	

Beliefs about benefits and obstacles by less-prepared individuals: Ten participants (71%) emphasized emotional-cognitive benefits related to crisis preparation, with a primary focus on the sense of security and independence (n=9), whereas 57% (n=8) cited emotional-cognitive drawbacks, i.e. fear (n=3) and stress (n=3). 79% (n=11) – more than in the well-prepared group – mentioned practical-material benefits, including improved crisis coping abilities (n=7) and increased flexibility in daily life (n=3). However, 93% (n=13) pointed to practical-material disadvantages, including the perceived effort (n=10), concerns about wastefulness (n=6) and fostering competition (n=3).

When it came to preconditions to preparedness, 79% (n=11) stressed the importance of knowledge and information, including its visibility (n=2) and recommendations (n=2). Around 21% (n=3) stated the value of a supportive network and practical assistance. Over 40% (n=6) mentioned a resource surplus of money, time, physical health, and available space.

Factors hindering access to behavior included lack of knowledge or information (57%, n=8), especially orientation (n=4) and lack of visibility (n=3), resource deficiency (36%, n=5), and a negative perspective on crisis preparation (50%, n=7). Notably, 57% (n=8) believed that nothing rendered crisis preparation impossible, attributing its feasibility to associated costs and effort (21%, n=3), while a third (n=4) discussed hypothetical factors like physical and psychological disabilities.

Table 7. Beliefs about Advantages, Group 1 = more prepared (n=16), Group 2 = less prepared (n=14)

	Group 1		Group 2	
	N	**%**	**N**	**%**
While preparing				
Emotional-cognitive	10	63%	10	71%
Sense of safety	*5*	*31%*	*9*	*64%*
Learning	*2*	*13%*	*1*	*7%*
Feeling competent	*3*	*19%*	-	
Practical-material	8	50%	11	79%
Better crisis management	*7*	*44%*	*7*	*50%*
Flexibility	*1*	*6%*	*3*	*21%*
Gains	-		*1*	*7%*
None	2	13%	-	
Preconditions				
Lack of knowledge/information	11	69%	11	79%
Topic visibility	*3*	*19%*	*2*	*14%*
Recommendations	*2*	*13%*	*2*	*14%*
Easy access	-		*1*	*7%*
Amount	*3*	*19%*	-	
Plan	*2*	*13%*	-	
Habit	*1*	*6%*	-	
Availability of resources	7	44%	6	43%
Money	*3*	*19%*	*3*	*21%*
Space	*3*	*19%*	*1*	*7%*
Time	*2*	*13%*	*3*	*21%*
Health	*1*	*6%*	*2*	*14%*
Network	5	31%	3	21%
Support	*2*	*13%*	*1*	*7%*
Exchange of experiences	*2*	*13%*	*1*	*7%*
Attitude	3	*19%*	1	7%
Positive gain expectancy	-		*1*	*7%*
Low effort expectancy	*3*	*19%*	-	
Enjoyment	*1*	*6%*	-	

Crisis Preparedness and Social Norm (RQ4): More Discussion and Less Disapproval in the Well-Prepared Group

Social norms were assessed by inquiring about the existence and style of discourse surrounding crisis preparedness in the social environment (see Table 8).

Perceived social norm by well-prepared individuals: Over two thirds mentioned discourse on the topic (69%, n=11%), dealing with practical strategies, information sharing, discussions on financial security and emotional coping. Half of the group (n=8) believe that their social environment prepares for disasters, noting measures towards energy independence, stockpiling, financial security, and efforts to stay informed. Regarding approval for crisis preparation behavior, around a third mentioned friends (38%, n=6) and family members (31%, n=5). Almost two thirds (63%, n=10) believed that nobody would disapprove of crisis preparedness.

Contrary, a quarter (25%, n=4) states that nobody in their environment approves of crisis preparedness. 25% (n=4) acknowledged that friends might not approve, and 19% (n=3) noted potential disapproval from many in the environment. 13% (n=2) mentioned disapproval from people in the local vicinity. A significant portion of the group felt that approval (19%, n=3) and disapproval (31%, n=5) depended on the extent of preparation.

Perceived social norm by less-prepared individuals: In this group the topic is less common in discussions (50%, n=7). Discourse surrounding stockpiling during the COVID-19 pandemic (n=2) and scenarios related to infrastructure failures in the context of the invasion of Russia in Ukraine (n=2) were mentioned, as well as the merit of insurances (n=2) and patient care (n=1).

57% (n=8) assume that their personal environment is not preparing, while 43% (n=6) thought that sufficient preparations were being made. Responses included mentions of stockpiling, financial preparations, survival training, and emotional readiness. In terms of approval of crisis preparedness, 57% in this group (n=8) stated that family members would approve, in about a third (36%, n=5), friends would approve, and 21% (n=3) mentioned their local environment who might profit would be supportive. 29% mentioned that approval and understanding depended on the measures being taken, indicating a strong negative perception of "excessive" crisis preparedness.

36% (n=5) of the group – less than in the well-prepared group – expected no one in their environment to disapprove of crisis preparedness. In contrast, three statements (21%) indicated that many in the environment would not approve, twice (14%) family members were mentioned.

Table 8. Perception of Social Norm, Group 1 = more prepared (n=16), Group 2 = less prepared (n=14)

	Group 1		Group 2	
	n	**%**	**n**	**%**
Discourse on crisis preparation behavior in social environment				
No	5	31%	7	50%
Yes	11	69%	7	50%
Practical	*7*	*44%*	*2*	14%
Resource-oriented	*2*	*13%*	*2*	14%
Financial	*2*	*13%*	*2*	14%
Emotional	*1*	*6%*	*1*	7%
Legal	-		*1*	7%
Forwarding information	*2*	*13%*	-	
Perceived crisis preparation behavior in social environment				
No	8	50%	8	57%
Yes	8	50%	6	43%
Supplies	*2*	*13%*	*3*	*21%*
Energy	*2*	*13%*	-	
Survival	-		*1*	*7%*
Emotional	-		*1*	*7%*
Financial	*2*	*13%*	*1*	*7%*
Information	*1*	*6%*	-	

Perceived acceptance of crisis preparation behavior in social environment								
	Approval		**Disapproval**		**Approval**		**Disapproval**	
Family members	5	31%	-		8	57%	2	14%
Friends	6	38%	4	25%	5	36%	1	7%
Local vicinity	1	6%	2	13%	3	21%	-	
Dimension-dependent	3	19%	5	31%	3	21%	4	29%
Many	-		3	19%	1	7%	3	21%
Nobody	4	25%	10	63%	-		5	36%

DISCUSSION

The quantitative results suggest that only one value is associated with crisis preparedness behavior: *benevolence*. With benevolence being defined as "preservation and enhancement of the welfare of people with whom one is in frequent personal contact" (Schwartz et al., 2000, p. 316), there might be an underlying belief that crisis preparedness contributes to the welfare of others, although this was not explicitly stated. Other values, such as *security*, which was found to be related to macro safety worry (Schwartz et al., 2000), showed no significant

correlation.

Given the importance of benevolence, questions concerning social, personal norms and beliefs could give insights into motivations. The results show that the more and less prepared individuals do not differ drastically in either of the dimensions. The sample described people who engage in crisis preparedness mainly as forward-thinking and reasonable, while a few negative associations pointed towards being egotistical or socially weak. There appeared to be a shared understanding that quite easily, crisis preparedness could become "too much", and people who prepare in this manner were labeled as "preppers" or "hoarders". Instead of emphasizing the positive social impact of crisis preparedness, e.g. by aiding emergency services, concerns about resource scarcity arose, with mentions of shortages during the COVID-19 pandemic, which was blamed on some people hoarding resources (Baddeley, 2020). Future research should investigate whether this notion persists when describing less resource-oriented crisis preparedness measures.

The qualitative content analysis indicated that the less-prepared group exhibited rather goal-oriented motivations, emphasizing a sense of security. In contrast, the well-prepared group valued the process itself, indicating competence and learning opportunities. They also demonstrated a more nuanced perspective considering preparation as a spectrum rather than a binary concept, resulting in a more practical focus on preparing a minimum supply of resources. These advantages may be due to more discourse about practical strategies in their social environment. The study found that over half of the participants perceived discussing crisis preparedness as unpleasant, associating it with fear, stress, and negative emotions. Many held the belief that it is time-consuming, and can lead to waste and competition. Further barriers to preparedness include a lack of differentiation between responsible, cooperative preparedness behavior and "hoarding". Many believe that more accessible information, recommendations, and emergency plans are needed.

From the findings, we can derive strategies for motivating people towards engaging in crisis preparedness. Firstly, given the relevance of benevolence and the negative image of "too much" preparedness, the social benefits of crisis preparedness should be stresses, especially the ways in which they positively affect one's immediate environment. A previous study used interactive nudges, which confronted users with the economic costs and benefits of (non-)preparedness and with the impact on their social environment. Positively framed statistical information, that asked users to guess costs, benefits and consequences of hazards and preparedness measures were found to be effective in increasing crisis preparedness (Haunschild et al., 2023), whereas the social nudges lead users to reconsider their behavior. In light of the identified contestation of crisis preparedness as a pro-social behavior, similar information could be used to stress the social benefits of preparedness, e.g. by confronting with the social consequences or reducing the perceived distance to hazards (Caraban et al., 2019). Research shows that interventions, such as eliciting reasonings for the relevance of a value, can lead to an increased perceived importance of that value (Russo et al., 2022). A future experiment should explore how elicitation of benevolence might affect attitudes towards crisis preparedness.

A second strategy could focus on showing the gaps and waiting times that are inherent in state-provided crisis management or in case of infrastructure failures (Reuter, 2014). This would be especially promising in countries with a state-oriented risk culture, which tend to rely overly on state institutions to manage crises (Reuter et al., 2019). While state institutions were hardly mentioned, there seems to be a common understanding that little individual preparedness is required, as predicted by risk culture theory. Future work should compare other risk cultures to see if attitudes change as the theory predicts.

Finally, in addition to offering more information on crisis preparedness and its effects, agencies should try to activate networks, encouraging individuals to discuss crisis preparedness in their immediate circles can facilitate easier adoption of the behavior. Integrating crisis preparedness into societal life is seen as a means to simplify access. Additionally, strategies from the Persuasive Systems Design model (Oinas-Kukkonen & Harjumaa, 2009), can be employed and potentially integrated into existing warning apps, offering rewards for engagement to make crisis preparedness more appealing.

The study is limited by its relatively small sample size, and future studies should seek to reproduce the connection between benevolence and crisis preparedness and possibly other values. While the median age of the sample (26-30) is younger than median age in Germany (45 years), research indicates that values (apart from conformity values, which tend to increase) remain substantially stable during adulthood (Vecchione et al., 2016). Despite efforts to stress the acceptability of all views, due to the focus of the study, participants may have perceived a desirability of crisis preparedness, introducing potential desirability bias. Additionally, the less-prepared group has lower levels in values overall, possibly influenced by the adapted questionnaire design tailored to crisis preparation behavior. As the items of the PVQ were adapted to crisis preparation behavior, this group may have been less able to identify with the values, which may explain the lower scores. A greater difference in values may have been achieved without adapting the questionnaire. Furthermore, the crisis preparedness score, relying on

official recommendations (BBK, 2019), focused on easily implementable measures. Future research may include measures that require more time, effort or financial investments, such as insurances or construction measures (Mol et al., 2021), to further differentiate people with a high intention and motivation to prepare.

CONCLUSION

This study has investigated connections between crisis preparedness and personal values, social and personal norms and beliefs. Differentiating more-prepared and less-prepared participants, the study has produced the following key findings:

- The quantitative analysis showed a significant correlation between greater preparedness and the personal value of *benevolence*.
- The qualitative thematic analysis of the interviews showed rather small differences regarding norms and beliefs between the two groups.
- A dominant social norm cannot be identified.
- A limited amount of crisis preparedness is believed to be helpful.
- A prominent theme in both groups is a negative view of "too much" preparation, leading to people being perceived as "hoarders", indicating that certain types of resource-oriented crisis preparedness are perceived as anti-social.

The study discusses strategies, such as confrontational nudges and value elicitation, for stressing the social benefits of crisis preparedness and triggering the value benevolence in this context. Recommendations for future research involve further examination of the impact of social norms and social structures, and investigating further countries characterized by different risk cultures (Reuter et al., 2019). Furthermore, the need for larger sample sizes, cluster analyses as in earlier research (Gerhold et al., 2019), and innovative approaches to promoting crisis preparedness behavior and behavior change are emphasized.

ACKNOWLEDGEMENTS

This research was funded by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE and by the LOEWE initiative (Hesse, Germany) within the emergenCITY centre. We thank Dr. Nina Gerber and Professor Verena Zimmermann for their advice and valuable feedback.

REFERENCES

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.

Appleby-Arnold, S. & Brockdorff, N. (2018). Culture and disaster risk management-synthesis of citizens' reactions and opinions during 6 Citizen Summits: Romania, Malta, Italy, Germany, Portugal and the Netherlands. CARISMAND report. https://www.um.edu.mt/library/oar/handle/123456789/53192

Appleby-Arnold, S., Brockdorff, N., Jakovljev, I., & Zdravković, S. (2021). Disaster preparedness and cultural factors: A comparative study in Romania and Malta. *Disasters*, *45*(3), 664–690. https://doi.org/10.1111/disa.12433

Baddeley, M. (2020). Hoarding in the age of COVID-19. *Journal of Behavioral Economics for Policy*, *4*(COVID-19 Special Issue), 69–75.

BBK. (2013). Kapazitäten der Bevölkerung bei einem Stromausfall. In *Empirische Untersuchung für das Bezugsgebiet Deutschland*. https://www.lkee.de/media/custom/2112_3209_1.PDF?1434529466

BBK. (2019). Katastrophenalarm. Ratgeber für Notfallvorsorge und richtiges Handeln in Notsituationen. In *Ratgeber für Notfallvorsorge und richtiges Handeln in Notsituationen* (7. Auflage). https://www.bbk.bund.de/SharedDocs/Downloads/DE/Mediathek/Publikationen/Buergerinformationen/Ratgeber/ratgeber-notfallvorsorge.pdf?__ blob=publicationFile&v=32

Bundeszentrale für politische Bildung. (2020). *Einkommen privater Haushalte*. https://www.bpb.de/kurz-knapp/zahlen-und-fakten/soziale-situation-in-deutschland/61754/einkommen-privater-haushalte/

Canlas, I. P., Karpudewan, M., & Mohamed Ali Khan, N. S. (2022). More Than Twenty Years of Value-Belief-Norm Theory of Environmentalism: What Has Been and Yet To Be Done? *Interdisciplinary Journal of Environmental and Science Education*, *18*(2), e2269. https://doi.org/10.21601/ijese/11801

Caraban, A., Karaponos, E., Goncalves, D. & Campos, P. (2019). 24 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. Conference on Human Factors in Computing Systems (*CHI19)*, https://doi.org/10.1145/3290605.3300733.

Cohen, J. (1988). *Statistical Power Analysis for Behavioral Sciences (2nd Ed)*. Lawrence Erlbaum Associates.

Cornia, A., Dressel, K., & Pfeil, P. (2016). Risk cultures and dominant approaches towards disasters in seven European countries. *Journal of Risk Research*, *19*(3), 288–304. https://doi.org/10.1080/13669877.2014.961520

Etkin, D. (1999). Risk transference and related trends: Driving forces towards more mega-disasters. *Environmental Hazards*, *1*(2), 69–75. https://doi.org/10.3763/ehaz.1999.0109

Federal Statistical Office. (2020). *Educational attainment of the population in Germany*. https://www.destatis.de/EN/Themes/Society-Environment/Education-Research-Culture/Educational-Level/Tables/educational-attainment-population-germany.html

Federal Statistical Office. (2023a). *Foreign population, 2015 to 2022 by selected citizenships*. https://www.destatis.de/EN/Themes/Society-Environment/Population/Migration-Integration/Tables/foreigner-citizenship-time-serie.html

Federal Statistical Office. (2023b). *Haushalte nach Haushaltsgrößen im Zeitvergleich*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Tabellen/1-3-privathaushalte-neuer-zeitvergleich.html

Fischer-Preßler, D., Bonaretti, D., & Fischbach, K. (2022). *A Protection-Motivation Perspective to Explain Intention to Use and Continue to Use Mobile Warning Systems*. *64*(2), 167–182. https://doi.org/10.1007/s12599-021-00704-0

Garschagen, M., & Sandholz, S. (2018). The role of minimum supply and social vulnerability assessment for governing critical infrastructure failure: Current gaps and future agenda. *Natural Hazards and Earth System Sciences*, *18*(4), 1233–1246. https://doi.org/10.5194/nhess-18-1233-2018

Garschagen, M., Sandholz, S., Wannewitz, M., Sabelfeld, R., & Krist, L. (2019). *Stand der Vorbereitung der Bevölkerung auf längere KRITIS-Ausfälle: Vorbereitung auf KRITIS-Ausfälle aus Endnutzer-Perspektive. Verortung von Verantwortung, Verwundbarkeit und Mindestversorgung*. https://kirmin.web.th-koeln.de/?page_id=707

Gerhold, L., Wahl, S., & Dombrowsky, W. R. (2019). Risk perception and emergency food preparedness in Germany. *International Journal of Disaster Risk Reduction*, *37*, 1–10. https://doi.org/10.1016/j.ijdrr.2019.101183

Goldstein, N. J., & Cialdini, R. B. (2007). Using social norms as a lever of social influence. In A. R. Pratkanis (Ed.), *Frontiers of social psychology. The science of social influence: Advances and future progress* (pp. 167–192). Psychology Press.

Grosz, M. P., Schwartz, S. H., & Lechner, C. M. (2021). The longitudinal interplay between personal values and subjective well-being: A registered report. *European Journal of Personality*, *35*(6), 881–897. https://doi.org/10.1177/08902070211012923

Haunschild, J., Pauli, S., & Reuter, C. (2023). Preparedness nudging for warning apps? A mixed-method study investigating popularity and effects of preparedness alerts in warning apps. *International Journal of Human-Computer Studies*, *172*. https://doi.org/10.1016/j.ijhcs.2023.102995

Honka, A., Helander, E., Pavel, M., Jimision, H., Mustonen, P., Korhonen, I., & Ermes, M. (2019). Exploring associations between the Self-Reported Values, Well-Being, and Health Behaviors of Finnish Citizens. *JMIR Mental Health*, *6*(4), 1–18.

Kim, S. H., & Seock, Y.-K. (2019). The roles of values and social norm on personal norms and pro-environmentally friendly apparel product purchasing behavior: The mediating role of personal norms. *Journal of Retailing and Consumer Services*, *51*, 83–90. https://doi.org/10.1016/j.jretconser.2019.05.023

Kohn, S., Eaton, J. L., Feroz, S., Bainbridge, A. A., Hoolachan, J., & Barnett, D. J. (2012). Personal disaster preparedness: An integrative review of the literature. *Disaster Medicine and Public Health Preparedness*, *6*(3), 217–231. https://doi.org/10.1001/dmp.2012.47

Latané, B., & Darley, J. M. (1976). *Help in a crisis: Bystander response to an emergency*. General Learning Press.

Mayring, P. (2015). *Qualitative Inhaltsanalyse*. Beltz.

Menski, U. (2016). *Neue Strategien der Ernährungsnotfallvorsorge. Ergebnisse aus dem Forschungsverbund NeuENV*. Freie Universität Berlin. https://doi.org/10.17169/refubium-21512

Mol, J. M., Botzen, W. J. W., Blasch, J. E., Kranzler, E. C., & Kunreuther, H. C. (2021). All by myself? Testing descriptive social norm-nudges to increase flood preparedness among homeowners. *Behavioural Public Policy*, 1–33. https://doi.org/10.1017/bpp.2021.17

Nohl, J., & Thiemecke, H. (1988). *Systematik zur Durchführung von Gefährdungsanalysen*. Fachverlag NW.

Nwafor, O., Singh, R., Collier, C., DeLeon, D., Osborne, J., & DeYoung, J. (2021). Effectiveness of nudges as a tool to promote adherence to guidelines in healthcare and their organizational implications: A systematic review. *Social Science & Medicine*, *28*, 1–13.

Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive Systems Design: Key Issue, Process Model, and

System Features. *Communications of the Association for Information Systems*, *24*(1), 485–500. https://doi.org/10.17705/1cais.02428

Reuter, C. (2014). Communication between Power Blackout and Mobile Network Overload: *International Journal of Information Systems for Crisis Response and Management*, *6*(2), 38–53. https://doi.org/10.4018/ijiscram.2014040103

Reuter, C., Kaufhold, M.-A., Schmid, S., Spielhofer, T., & Hahne, A. S. (2019). The impact of risk cultures: Citizens' perception of social media use in emergencies across Europe. *Technological Forecasting and Social Change*, *148*, 1–17. https://doi.org/10.1016/j.techfore.2019.119724

Reynolds, K. J., Subašić, E., & Tindall, K. (2015). The Problem of Behaviour Change: From Social Norms to an Ingroup Focus. *Social and Personality Psychology Compass*, *9*(1), 45–56. https://doi.org/10.1111/spc3.12155

Russo, C., Danioni, F., Zagrean, I, & Barni, D. (2022). Changing Personal Values through Value-Manipulation Tasks: A Systematic Literature Review Based on Schwartz's Theory of Basic Human Values. *European Journal of Investigation in Health, Psychology and Education*, *12*(7), 692-715.

Schwartz, S. H. (1977). Normative Influences on Altruism. *Advances in Experimental Social Psychology*, *10*, 221–279. https://doi.org/10.1016/S0065-2601(08)60358-5

Schwartz, S. H. (1992). Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. *Advances in Experimental Social Psychology*, *25*, 1–65. https://doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S. H. (2016). Basic individual values: Sources and consequences. In D. Sander & T. Brosch (Eds.), *Handbook of value* (Vol. 4, pp. 1–23). Oxford University Press.

Schwartz, S. H., & Cieciuch, J. (2022). Measuring the Refined Theory of Individual Values in 49 Cultural Groups: Psychometrics of the Revised Portrait Value Questionnaire. *Assessment*, *29*(5), 1005–1019. https://doi.org/10.1177/1073191121998760

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., Demirutku, K., Dirilen-Gumus, O., & Konty, M. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, *103*(4), 663–688. https://doi.org/10.1037/a0029393

Schwartz, S. H., Sagiv, L., & Boehnke, K. (2000). Worries and Values. *Journal of Personality*, *68*(2), 309–346. https://doi.org/10.1111/1467-6494.00099

Statista. (2023a). *Population in Germany: Average age 2011-2022*. https://www.statista.com/statistics/1127805/population-average-age-germany/

Statista. (2023b). *Population of Germany 1990-2022, by gender*. https://www.statista.com/statistics/454338/population-by-gender-germany/

Steg, L., De Groot, J. I. M., Dreijerink, L., Abrahamse, W., & Siero, F. (2011). General Antecedents of Personal Norms, Policy Acceptability, and Intentions: The Role of Values, Worldviews, and Environmental Concern. *Society & Natural Resources*, *24*(4), 349–367. https://doi.org/10.1080/08941920903214116

Steg, L., Dreijerink, L., & Abrahamse, W. (2005). Factors influencing the acceptability of energy policies: A test of VBN theory. *Journal of Environmental Psychology*, *25*(4), 415–425. https://doi.org/10.1016/j.jenvp.2005.08.003

Stern, P. C. (2000). New Environmental Theories: Toward a Coherent Theory of Environmentally Significant Behavior. *Journal of Social Issues*, *56*(3), 407–424. https://doi.org/10.1111/0022-4537.00175

Thøgersen, J., & Ölander, F. (2002). Human values and the emergence of a sustainable consumption pattern: A panel study. *Journal of Economic Psychology*, *23*(5), 605–630. https://doi.org/10.1016/S0167-4870(02)00120-4

UNDRR. (2016). *Report of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction* (pp. 1–41). United Nations Office for Disaster Risk Reduction. https://www.undrr.org/publication/report-open-ended-intergovernmental-expert-working-group-indicators-and-terminology

UP KRITIS. (2014). *Öffentlich-Private Partnerschaft zum Schutz Kritischer Infrastrukturen. Grundlagen und Ziele* (pp. 1–54). BSI. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KRITIS/UPK/upk-grundlagen-ziele.pdf? blob=publicationFile&v=6

Vecchione, M., Schwartz, S., Alessandri, G., Döring, A. K., Castellani, V., & Caprara, M. G. (2016). Stability and change of basic personal values in early adulthood: An 8-year longitudinal study. *Journal of Research in Personality*, *63*, 111–122. https://doi.org/10.1016/j.jrp.2016.06.002

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*(5), 806–820. https://doi.org/10.1037/0022-3514.39.5.806

Worldbank. (2023). *Urban population - Germany*. https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=DE

Engaging the Public in Forest Fire Awareness Through a Mobile Application

Emil Gatial
Institute of Informatics, Slovak Academy of Sciences, Slovakia
emil.gatial@savba.sk

Zoltán Balogh
Institute of Informatics, Slovak Academy of Sciences, Slovakia
balogh@savba.sk

Sepideh Hassankhani Dolatabadi
Institute of Informatics, Slovak Academy of Sciences, Slovakia
sepideh.ui@savba.sk

Mariana Soledad Molina
Massive Dynamic, Sweden
marsmolina20@gmail.com

Timo Kasig
Massive Dynamic, Sweden
timo.kasig@gmail.com

Kayvan Yousefi Mojir
Faculty of Librarianship, Information, Education and IT, University of Borås, Sweden
kayvan.yousefi_mojir@hb.se

Marcin Wawrzyniak
ITTI Sp. z o.o. , Poland
marcin.wawrzyniak@itti.com.pl

Dominika Grunwald
ITTI Sp. z o.o. , Poland
dominika.grunwald@itti.com.pl

Marcin Przybyszewski
ITTI Sp. z o.o. , Poland
mprzybysz@itti.com.pl

ABSTRACT

The widespread use of mobile devices has resulted in remarkable advancements in their capacity to offer and disseminate online information virtually anywhere. Users are increasingly open to installing applications that furnish valuable information to address specific situations. This paper describes joint efforts in research and development of a mobile application, called Citizen Engagement App (CEA), and related methods for engaging the public in forest fire danger awareness, data collection and presentation services. The primary objective is to provide insights into the developed methods and technologies employed. The methods are separated into two domains, namely the domain of citizen engagement approaches to educate people about forest fire prevention, and the domain of mobile application development employed for disseminating knowledge and reporting on hazardous situations in the forests. The evaluation of the mobile application took place both in the pilot demonstrations and meetings with the public to pinpoint any gaps and challenges.

Keywords

Wildfire, forest fire prevention, citizen engagement, forest fire reporting, mobile application.

INTRODUCTION

The widespread use of mobile devices has resulted in remarkable advancements in their capacity to offer and disseminate online information virtually anywhere. The users are increasingly open to installing applications that furnish valuable information to address specific situations. Even though users' willingness at data collection attendance is considered to be low (Wenz, 2023), as the use of mobile technologies for personal purposes increases, including among older groups in the population, participation in survey-based activities using these technologies is likely to increase (Jäckle, 2019). As a general rule, globally only around 4% of all forest fires have natural causes such as lightning. In all other cases, humans are responsible for the fires - be it directly or indirectly, deliberately or due to carelessness (Hirschberger, 2016). Aforementioned facts about the increasing use of mobile applications as a communication information channel and that major cause of wildfires is a human activity was the main drivers for creation of mobile application for forest fire prevention. Since the preventive actions help to decrease the forest fire causes, the sharing the information about the existing forest fires may help to put the fire under control or have better situational overview. This paper also addresses the concern of people's willingness to install and actively use the application especially during the wilderness adventures. The presented work is being developed in scope of the SILVANUS project[1].

In the next chapter, the paper provides a related work done in the domain of citizen engagement for preventing forest fires using campaigns and the use of mobile applications for forest fire reporting/monitoring. The focuses on the design, development, utilization, and acceptance of a mobile application aimed at providing the useful information on wildlife safety, the importance of wildfire prevention and sharing information about hazardous situations in the forests. The following chapters are describing the methods raising public awareness about the importance of the forest fire prevention through various channels as social media, workshops, exhibitions and especially mobile application. The next chapter focuses on mobile application design with regard on user experience, followed by the fire reporting module design and implementation of backend services. Next chapters are devoted to providing overview on the information collection and distribution framework directly used by the mobile application and the main dashboard description. The last chapters of this paper discuss the prototype application evaluation and identified challenges along with the future work addressing the challenges.

RELATED WORK

The chapter summarizes scientific works related to mobile applications and citizen engagement for fire prevention and fire reporting/warning, and also reviews the existing mobile applications in this domain. The topics covered in reviewed works include the mobile apps, analysis of emergency notification apps, integration of mobile apps into emergency communication frameworks, examination of mobile crowdsensing technologies in emergencies, and recent developments in wildfire monitoring. The mobile application EscapeWildFire benefits was described (Kamilaris, 2023) and evaluated through simulations of wildfires in Cyprus and the 2011 Amarillo, Texas wildfire to aid in real-time wildfire evacuation. The work overviewed the app's wildfire-specific design requirements. The set of the most relevant emergency notification mobile applications were reviewed and analyzed (Romano, 2016) according to the shared data types and forms of communication with citizens. One of the conclusions states that people prefer to use multi-media to write a text or fill in the forms for notifying emergencies. In the scope of PEMEA project (Pemea, 2020), twelve mobile emergency applications are integrated to the communication framework defined by ETSI (TS 103 478) standard providing citizens with fast and intuitive access to emergency services. The work (Cicek, 2023) provides a comprehensive overview of the mobile crowdsensing technologies in emergency situations as well as research and development challenges. The work concludes with a finding that mobile applications and social media prevail in hazard/risk detection, evacuation/mapping, information exchange, situational awareness and data fusion. There was a recent effort in design and develop mobile application (Teguh, 2021) to control and monitor peat fires in the Central Kalimantan (Indonesia) thus providing firefighters with information support. The requirements study (Majlingova, 2022) enlisted specific requirements for mobile application relevant to wildfire management. The following citizen requirements were identified: notification about the fire occurrence in citizens' vicinity; ability to notify the fire location to the fire and rescue service; ability to notify the forest management services on human negligence; and ability to report a suspect fire with geo-location information, photos, and a brief description. The Smoke Sense project by the US EPA (Environmental Protection Agency) utilizes a mobile app to involve citizens at different tiers to mitigate the public health effects of wildland fire smoke, drawing motivation from personal, professional, and community contexts and collecting individual reports of smoke exposure. It underscores the necessity of aligning citizen science initiatives like Smoke Sense with partner organizations to tackle intricate problem areas effectively (Hano, 2020). Although forest fire reports can also be obtained from social media through topic extraction method (Messaoudi, 2022), the

[1] SILVANUS Horizon 2020 Green Deal project https://silvanus-project.eu/

recent evolution of security and privacy getting these sources more restricted.

Apart from the scientific research, the table (Table 1) summarized the unique features of the most relevant mobile applications in wildfire monitoring and warning. The main purpose of the reviewed applications is to offer real-time updates on wildfire locations and status to keep users informed about the latest developments. The information is typically placed on the interactive maps where users can view fire incidents, evacuation routes, hotspots, and other relevant information. Another common functionality is receiving alerts and notifications about nearby wildfires, weather conditions, evacuation orders, and other critical updates. The applications commonly use satellite data for wildfire warning for global coverage or rely on fire department information sources used for local state coverage, specifically only for regions encompassing United States, Canada and Australia.

Table 1. Existing wildfire emergency applications.

	OS/Plan	Region	Summary of unique functionalities
Watch Duty	Android/iOS, Free	US	Relies on real people, including trained professionals like firefighters and dispatchers, to provide real-time information verified by humans rather than automated systems.
Fireguard Wildfire Tracker	Android, Ads	World	Global monitoring of wildfires using satellite-based activity tracking and provides near real-time updates on fire locations without requiring users to log in, but provides inaccurate/unverified alerts.
Fires Near Me Australia	Android/iOS, Free	AU	Focuses on monitoring nationwide bushfire threats in Australia and categorizes alerts into Advice, Watch, and Act levels based on the Australian Warnings System.
Wildfire Tracker	iOS, Ads/Paid	US, Canada	Provides detailed weather forecasts, air quality data, and an interactive map showing wildfire hotspots, along with push notifications and global updates.
Frontline Wildfire Tracker	Android/Free	US, Canada	Comprehensive situational awareness including event notifications and preparation checklists.
Fire Atlas	iOS, Ads/Paid		Includes additional maps from the Bureau of Land Management and the US Forest Service, with features such as offline mapping, thermal hotspots, and AccuWeather summaries.
Firespot: Wildfire app	iOS, Ads/Paid	World	Offers real-time global wildfire information with satellite and for US/Canada agency data, including perimeters, aircraft tracking, and customized alerts for subscribers.
Bushfire.io	Android/iOS, Free/Paid	AU	Provides immediate alerts, up-to-date maps, and collaboration with local authorities for verified data, along with features like emergency aircraft positions and wind forecasts.
Australian Fires	iOS, Free	AU	Aggregates data from multiple government agencies covering fire incidents across Australia, emphasizing the importance of staying informed through official channels.

CITIZEN ENGAGEMENT APPROACH

In our forest fire prevention efforts, we have implemented various strategies to engage citizens and raise awareness about the importance of wildfire prevention. One of our key initiatives involves the creation and distribution of informative posters in communities at risk of forest fires. These posters contain essential information about fire prevention measures, such as safe campfire practices, proper disposal of cigarette butts, and the importance of reporting any signs of smoke or fire promptly. By combining traditional outreach methods like posters with

modern technological solutions like the CEA application, we aim to foster a collaborative approach to forest fire prevention, involving communities in safeguarding our precious natural resources.

Overall Approach Overview

We view citizen engagement as a result of bidirectional communication, involving citizens, organizations, and authorities in wildfire management (Mojir, 2023). Media effects theory was incorporated into the definition of citizen engagement, viewing media as a means for audiences to acquire knowledge, attitudes, emotions, and behavior change (Bandura, 2001; Potter, 2012). The critical focus of the Citizen Engagement Programme (CEP) lies in determining effective communication means and channels. These tools are essential for informing citizens, aiding in wildfire preparedness, prevention, and protection efforts. Additionally, the program emphasizes active citizen participation in shaping wildfire prevention policies, supporting first responders, and engaging in post-disaster recovery activities. Figure 1 outlines simplified effects desired from citizen engagement, marked as "raising awareness," "changing attitude," and the more intricate "changing behavior" in the context of wildfire management. The figure also includes "Type of activity" aligns with these effects, including informing and educating for awareness, fostering safe practices for attitude change, and organizing assistance for behavioral change (Mojir, 2023). The desired effects are relevant in all phases of wildfire management including prevention, response and recovery.

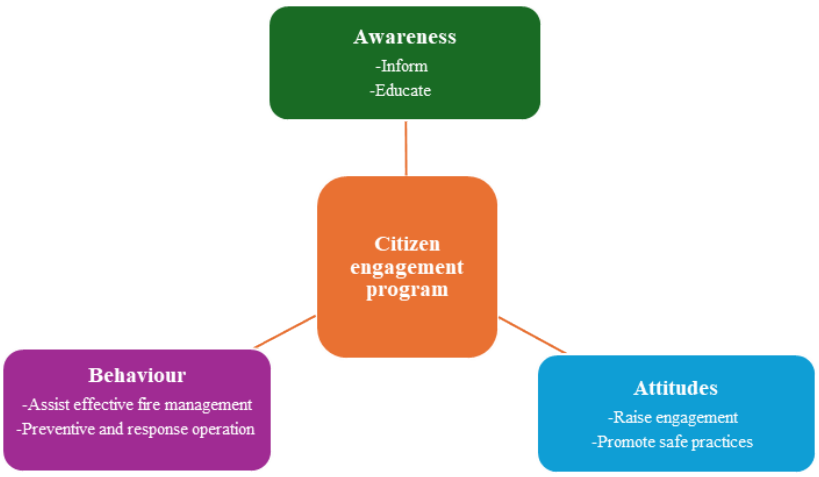


Figure 1. Simplified model of desired effects concerning citizen engagement program in wildfire management.

Examples of Content and Achievements

As part of the CEP, social media campaigns and a poster exhibition in different cities were held, involving the creation of impactful messages with visuals to prompt immediate realization or comprehension. The team, now including members with communication and graphic design expertise, manages these campaigns on social media. Additionally, work is underway to develop an online citizen engagement course based on microlearning, set to be published soon for citizens' use in various settings such as workshops, training sessions, or as a self-study course. The online course serves as a valuable tool for engaging citizens in wildfire prevention, response, and recovery. It provides educational content, empowers citizens with knowledge and skills, raises awareness, and encourages participation. A mobile application was also designed, primarily to enable citizens to assist in wildfire management processes. The mobile application will be explained in more detail. Figure 2 provides examples of published posters on social media.



Figure 2. Examples of posters designed for social media and exhibitions

CITIZEN ENGAGEMENT MOBILE APPLICATION

The primary objective of the mobile application is to provide a comprehensive guide for users moving through forested areas. The app's aim is to equip the users with essential information and guidance on how to effectively handle unexpected situations that may arise during their outdoor activities. Users may access up-to-date information about the local environment to make them sure that they can be well-prepared for expected conditions. This includes details about potential hazards and specific guidelines to mitigate risks. Whether it's dealing with common issues like sunburns and injuries or accessing region-specific information, the application strives to enhance the overall experience and safety of users exploring nature.

In addition, the application encourages users to share valuable information to the local community by reporting and sharing observations related to potentially hazardous events like forest fire or wild animals in the forest areas, which contributes to the collective safety. Through a user-friendly interface, the app seamlessly integrates essential features, such as an emergency hotline, first aid tips, and location-based services. The goal is to empower users to make informed decisions while promoting a sense of responsibility towards the environment and the local community. By combining practical guidance with community engagement, the mobile application strives to make forest exploration not only an enjoyable experience but also a safer and more sustainable one for everyone involved.

Design and Navigation in the App

The Citizen Engagement App (referenced as the App) is developed to enhance user experience and accessibility, every feature or module is designed to facilitate smoother navigation and interaction for the users.

In general, the App enables a user to navigate using bottom bar to show main or home, map and configuration panels. This navigation provides users with quick access to the modules withing the app, allowing seamless switching between the different panels. The App's central panel provides a home screen presenting a comprehensive array of features, including informative news updates, detailed guidelines covering a variety of topics relevant to forest navigation, practical tips to enhance user experiences, and engaging quizzes designed to educate and entertain users as they explore forest environments. Every feature is presented to users with a visually appealing and informative card. Each card features an image and a concise description along with intuitive navigation option. This design aims to improve user engagement and streamline access to relevant content (Figure 3).

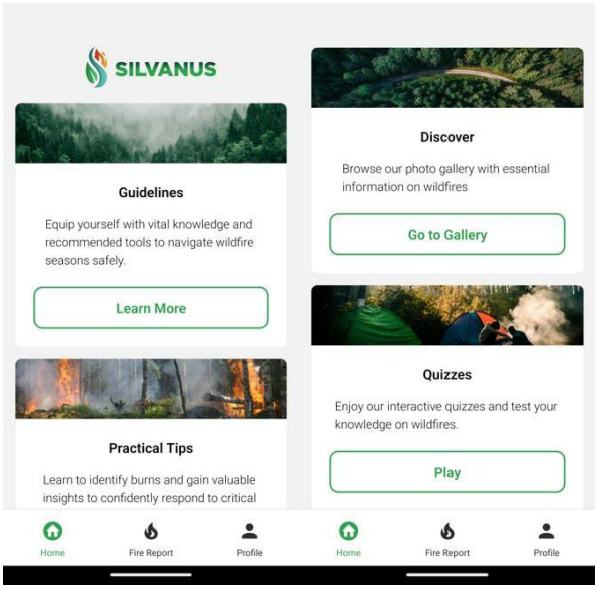


Figure 3. Screenshots of the home screen, showcasing the cards for the different modules.

Withing the configuration or profile panel there is a list of three clickable cards providing users with access to different functionalities. First, the language settings where users are presented with a dropdown menu offering a selection of languages. This feature enables users to customize the app interface and content according to their preferences. Enhancing accessibility and usability. Second, a Privacy Policy that redirects user to the documentation of the app. Third, the Q&A card provides users with access to frequently asked questions and additional information about the app and its underlying project (Figure 4).

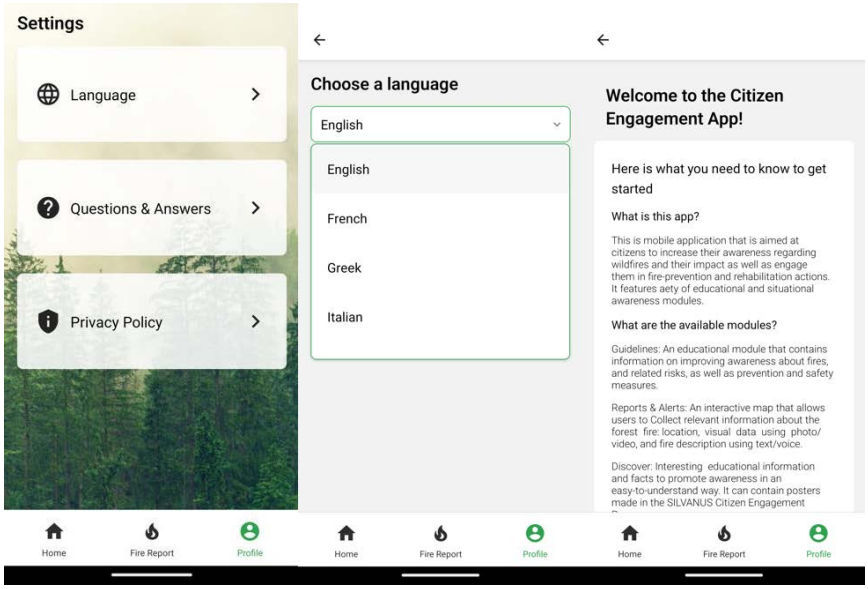


Figure 4. Profile panel. From left to right: main profile menu. Language configuration. Questions and answers information.

The map panel unwraps into the separate module used for reporting and receiving information using subscribed information channel mechanism (described in chapter Fire Reporting Module). The App's design feature is to display relevant information incorporated into the map. While focusing on the specific functionality of fire reporting, the information channels may collect information about different topics, which do not necessarily have to relate to the fire reporting.

Information and Guidance Modules

The guidelines module provides users with essential information and instructions related to wildfire prevention and safety measures. With a focus on empowering citizens residing in forested areas, this module offers

comprehensive guidance on various scenarios, including appropriate actions to take upon encountering a wildfire and post-fire instructions for safety and recovery. Upon accessing the Guidelines module, users are presented with a selection of eight topics covering a range of pertinent subjects. Navigating within each topic reveals a swipe mechanism, allowing users to seamlessly browse through informative cards that encapsulate key guidelines. Each card presents concise and actionable advice, with up to three guidelines displayed per swipe. For users seeking detailed information, a prompt button is available to access the complete list of guidelines within the selected topic (Figure 5).

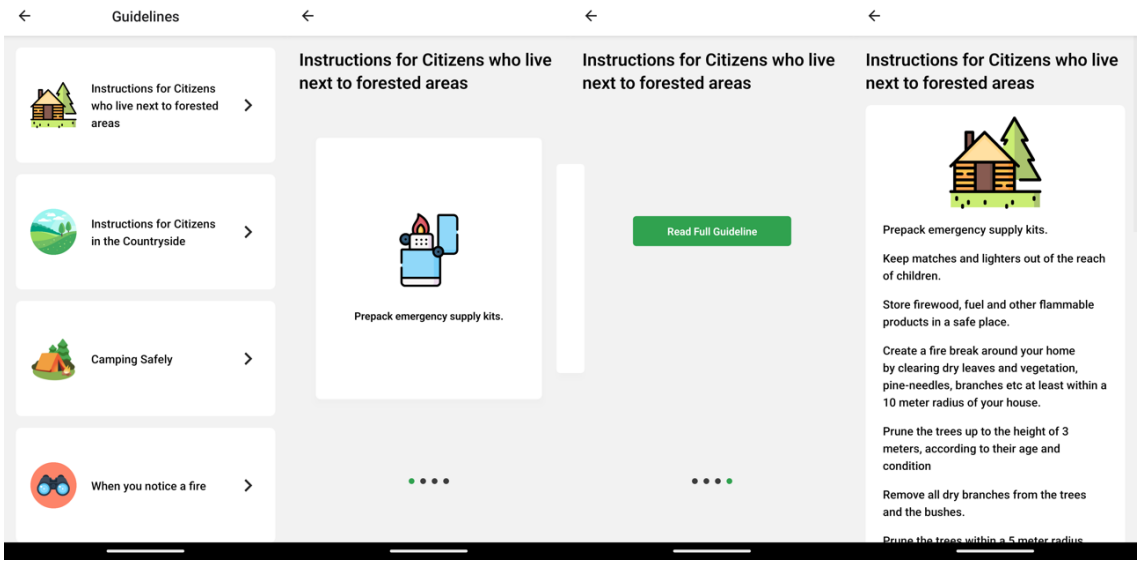


Figure 5. Screens for the guideline module. From left to right navigation of the screens.

In addition to the Guidelines module, the Citizen Engagement App features a Practical Tips section tailored to provide users with actionable insights for handling various emergency situations commonly encountered in forest environments. This module offers practical advice on topics such as first aid treatments for minor injuries, recognition of heatstroke symptoms, and understanding the causes of burns. Similar to the Guidelines module, the Practical Tips section consists of seven distinct titles, each focusing on a specific aspect of wilderness safety and emergency preparedness. Upon selecting a title, users are presented with a series of swipeable cards containing detailed tips and recommendations relevant to the chosen topic (Figure 6).

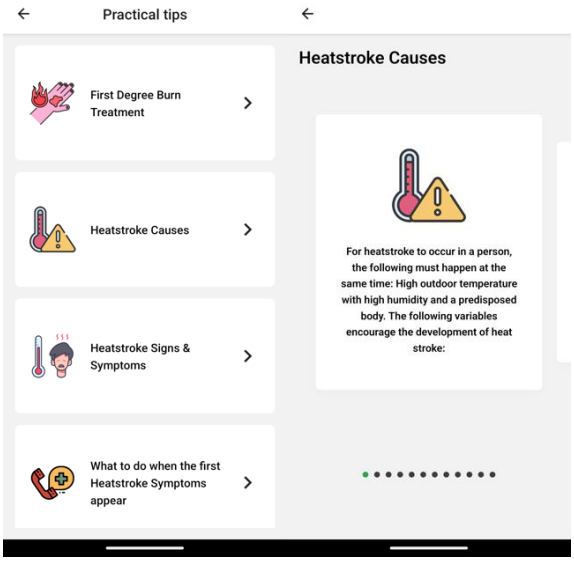


Figure 6. The screens of practical tips module. Left menu with titles. Right swipe screen.

The discover panel is designed to resemble a gallery, showcasing a range of content related to forest safety and wildfire prevention. Aiming to enrich user exploration and understanding of key topics through visual storytelling.

Upon accessing the Discover module, users are greeted with a selection of four distinct titles, each highlighting a specific theme relevant to forest safety and emergency planning. These titles include “Effects of Wildfires”, “Protect Our Home from Wildfires”, “Making a Family Emergency Plan” and “Causes of Wildfires.” Under each title, users encounter a horizontal scrolling gallery of images, providing a glimpse into the associated topic’s visual narrative. Clicking on an individual image opens it in full-screen mode, allowing users to immerse themselves in its details. From here, users can swipe horizontally to navigate through the entire collection of images within the selected title, enabling a seamless exploration experience (Figure 7).

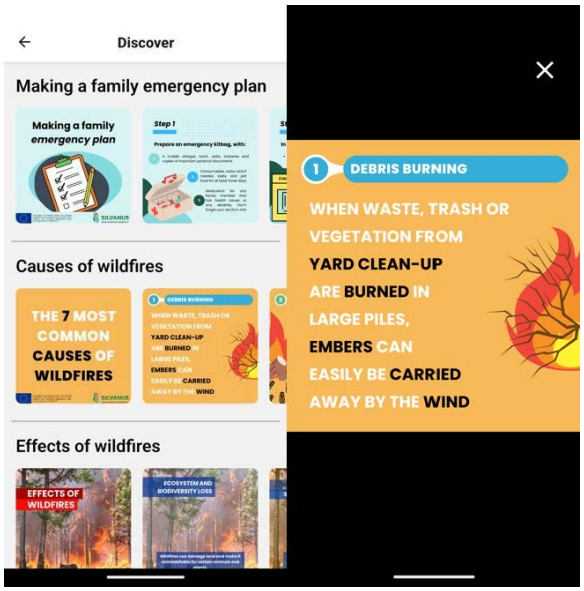


Figure 7. Screenshots of discover module. Left image gallery of photos. Right image full-screen mode.

The quizzes panel offers users an interactive and engaging way to test their knowledge on wildfire safety and emergency response procedures. With two quizzes currently available, the “Burns Quiz” and the “Heatstroke Quiz”, this module provides users with an opportunity to assess their understanding of critical topics related to wilderness safety. Upon selecting a quiz title, users are presented with a visually intuitive interface featuring a progress bar indicating the number of questions completed out of the total. Each quiz consists of multiple-choice questions designed to challenge users’ comprehension of key concepts and best practices. Upon completing the quiz, users receive a summary of their performance. If more than half of the answers were correct, users are congratulated with a message indicating their success and encouragement to continue their engagement with the app’s content. In contrast, if more than half of the answers were incorrect, users are provided with constructive feedback and encouraged to explore other modules within the app to improve their knowledge and future quiz scores. This iterative approach encourages continuous learning and reinforces key safety principles in forest environments (Figure 8).

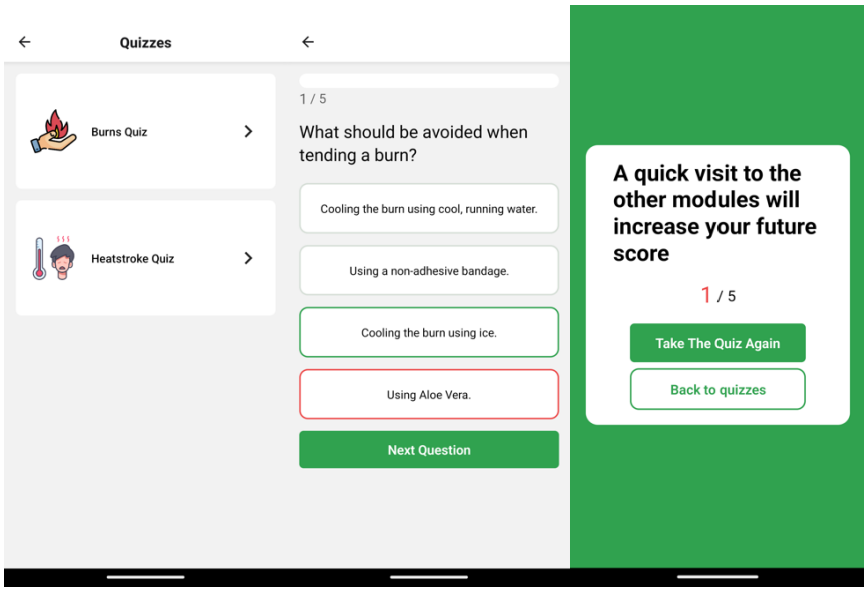


Figure 8. Quizzes module. Navigation through one quiz.

Fire Reporting Module

The Fire Reporting module is designed to provide valuable information on the main map (the first picture in Figure 9) while also facilitating the sharing of information across various information channels. The map interface shows a zoomable vector map featuring the user's current location. Positioned at the bottom right corner of the map is the fire reporting icon, which, when selected, places a fire icon at the center of the map. Through a long touch gesture, the user can update the fire icon location to accurately reflect the incident's location. Upon selecting the fire reporting icon once again, the information collection form is displayed (the second picture in Figure 9). Here, users are provided with the opportunity to select images from their smartphone's gallery or capture new photos using the device's camera. After selecting at least one thumbnail of the chosen picture and composing a brief message, users can transmit the data into information poll by clicking the submit button. The data transmission process is described in detail in the subsequent section, outlining the various parameters and mechanisms involved in the communication process.

The Channel screen (the third picture in Figure 9) serves as a hub for managing available information channels via services provided by the EmerPoll system. Within this screen, users can search among the available channels and based on their preferences, subscribe or unsubscribe to/from any channel. Subscribing to an information channel triggers the download of the corresponding poll template, which is then added to the user's active channels. Each poll template within the subscribed channels contains detailed instructions on how to construct the visual form, including visual icons and form data, as well as guidance on assembling the data packet for poll replies. Subscribed channels also influence the icon set displayed on the main map, but the behavior of the information reporting process remains the same as described for fire reporting albeit with different objectives.

The Message screen (the last picture in Figure 9) provides users with a list of messages sent or received. Here, users can review the details of each message apart of the message related location.

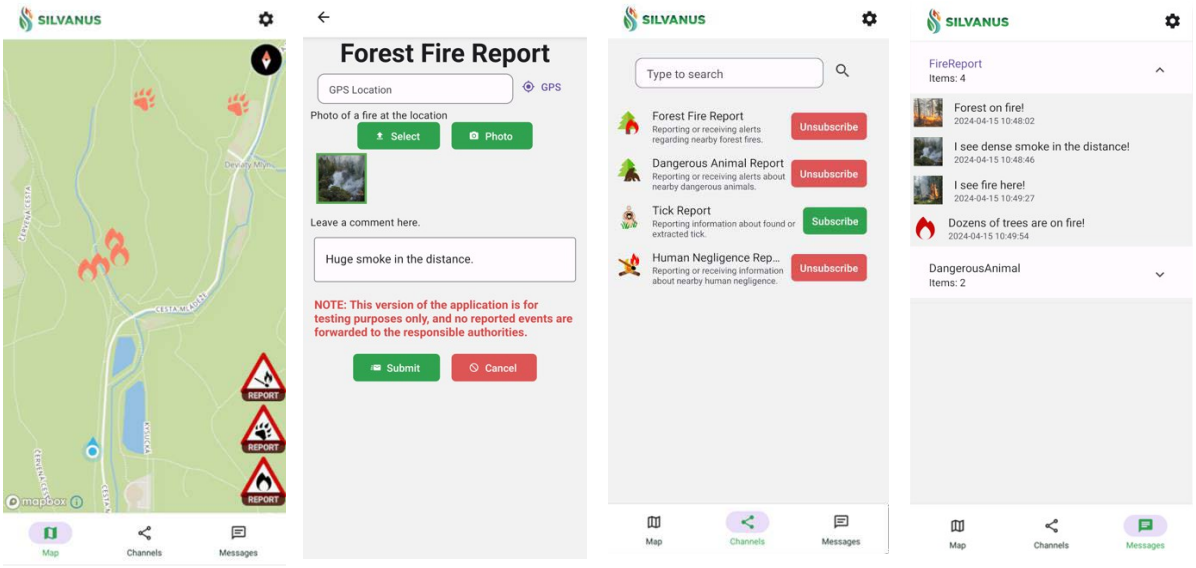


Figure 9. The screens of Fire Reporting module. From left: main map used for localizing the event, reporting form, channel management screen and message screen.

Implementation of Content and Backend Services

Implementation of mobile application based on the React Native and Expo frameworks, which allow development and deployment of mobile applications for both Android and iOS systems.

The backend services supporting the fire reporting module (depicted in Figure 10) are designed to prioritize scalability and straightforward deployment processes. Central to this infrastructure is the MQTT gateway, secured over an SSL layer facilitated by the EMQX server. This gateway serves as the primary communication channel, enabling the application to facilitate the transmission of polling messages of various channels with EmerPoll services and the user community registered to specific channels via MQTT topics. The polling message payload comprises of JSON-encoded data, primarily consisting of form data, accompanied by polling details, user information, and channel metadata. For immediate message overview, thumbnail images are included into the messages. File Storage service is employed to facilitate large binary data transfer, typically pictures, audio and video recordings captured by smartphones. This data is encoded into ‘multipart/form-data‘ format and transferred by secure HTTPS protocol. Uploaded binary data become accessible by the Nginx Web Server via links provided within the metadata of MQTT messages.

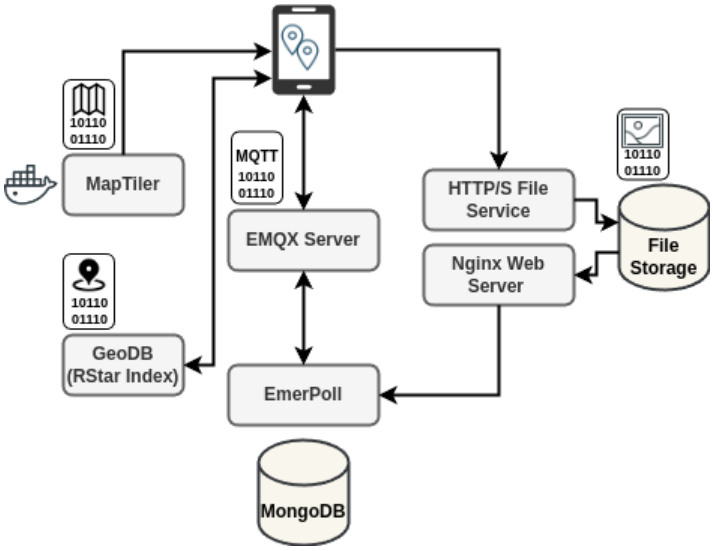


Figure 10. Fire reporting module backed services and data flow schema.

The CEA mobile application utilizes vector map tiles sourced from a containerized MapTiler server to render the Map screen. Geo-spatial information embedded within overlaid data is stored and queried using the R-star Tree

method, optionally supported by a key-value store like Leveled DB. This approach ensures efficient data management and seamless integration of spatial information within the application's interface.

INFORMATION COLLECTION AND PRESENTATION

Information Collection and Distribution Methods

Information collection method is based on the principles of Polls, Channels and Templates as described in (Balogh, 2016) and later extended by the information sharing protocols proposed in (Balogh, 2023). Actual integration with the CEA application follows the same principles that allow users to search and subscribe to information channels even other than for a wildfire warning and reporting as described in this paper. The operations such as information channel search use dedicated MQTT topic mimicking request-response communication. Information channel management actions use the MQTT topic infrastructure to facilitate the subscription process. EmerPoll utilizes the external services of the Storage Abstraction Layer (SAL), a component developed within the SILVANUS project, to facilitate the transfer of data to and from shared storage resources. SAL empowers EmerPoll by providing notifications regarding the availability of specific data, enabling seamless access for processing. This asynchronous capability allows data to be both sent and received efficiently, enhancing the overall functionality of EmerPoll.

Dashboard Data Integration Methods

Dashboard is a web-based, graphical interface created for SILVANUS platform, which is the main access point for the user to obtain and interact with outcomes of data produced by SILVANUS' various modules. Dashboard consists of layers representing the current status of chosen area of operation on the interactive map. Each layer is dedicated to individual types of data and events produced by a single User Product.

Data from EmerPoll is aggregated on the layer called Mobile Subscribers, representing emergency events reported by users in an area where occurrence of fire or smoke has been spotted. Events are visualized as a connected pair of pins representing the position of a user along with the position where the fire is taking place from the perspective of the user. Each of the singular events can be further inspected, presenting more details provided by the user. Embedded information is presented in the menu on the side, which contains exact coordinates to the location of the event, date, time, picture, description and status. Both description and picture are optional, depending on the action of the user. Example of the Mobile Subscribers layer from Dashboard is shown by Figure 11.

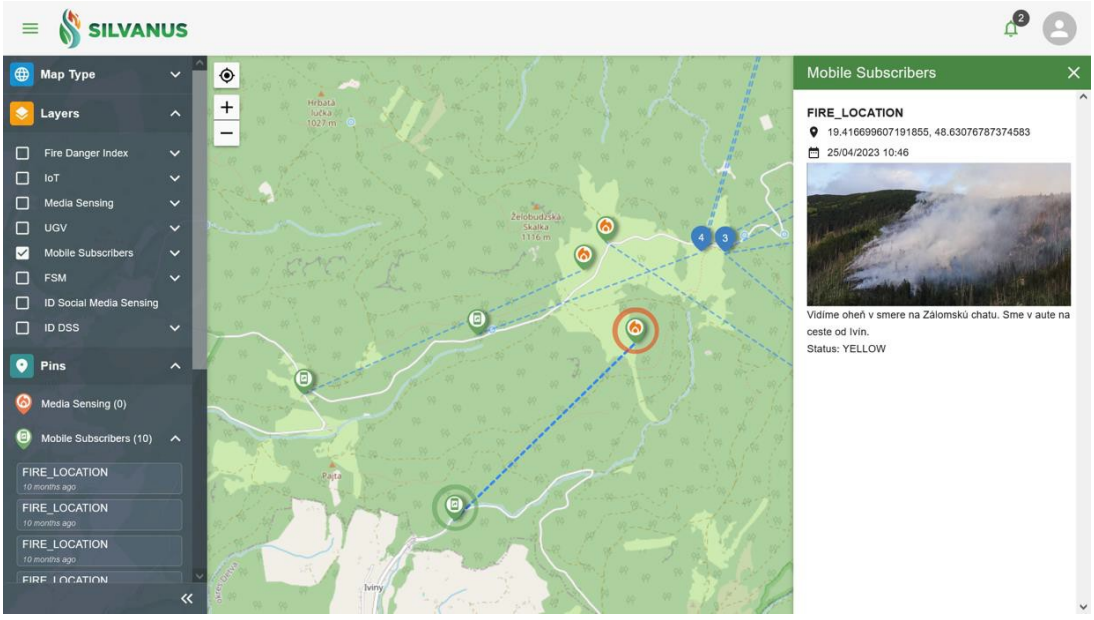


Figure 11. Mobile subscribers Dashboard layer.

The data for the Mobile Subscribers layer is available both live and archive, through the adequate message queue in the SILVANUS Storage Abstraction Layer.

Upon recognizing the danger in the area, Dashboard allows the operator to communicate directly with EmerPoll,

requesting a Poll to send a warning message to the subscribers of chosen channel.

EVALUATION AND CHALLENGES

The concept of the mobile application was first demonstrated during the Slovak pilot exercise located in Poľana region in the central Slovakia. The CEA was presented as a part of information collection framework called EmerPoll used for crowdsourced data collection and aggregation. The fire reporting data (photos and notes) were displayed on the interactive map of Dashboard Layer. The concept of pushing the wildfire warnings to smartphones was also demonstrated. The second evaluation of the technology was performed in autumn 2023 in Ostrava region located in north-east part of Czech Republic, where developed technologies was deployed in the remote command and control vehicle used to evaluate fire reporting functionality by CEA application during the fire fighters exercise. During the first evaluation phase, the coverage of mobile data connection and even slow transmission data rates was found as a limiting factor of the CEA application. To mitigate the factor of mobile data network intermittent connection along with slow data rate, the CEA adopted the off-line message acquisition using the MQTT protocol (e.g. messages are delivered as soon as the mobile data recovers) and also transmission of reported photo overview (picture thumbnails) in a single MQTT message, while full resolution picture is transferred using HTTPS protocol independently. The mobile data network coverage is panned to put as an overlay into the CEA interactive map to help users to find the places with mobile data connection in the forest area. During the second evaluation phase, the challenge with reporting the precise location was identified, where a user of mobile application may struggle to put a marker on interactive map in a correct direction. As a remedy to this situation, the virtual compass will be put into the interactive map along the distance estimation of the user from places marker for reported fire.

FUTURE WORK

The plans for future work were guided by the discussions with the potential users and first responders during the meetings and during the pilot exercises, as mentioned in the previous chapter. Specifically, the CEA is planned to be extended with the virtual compass and mobile data coverage map. The CEA application will be extended with customizable alerts that notify users of fire hazard events near a user-defined location. CEA is being prepared to acquire and display data containing evacuation routes and wildfire area temporal evolution, which will be provided by SILVANUS services. The Profile or Configuration panel's current functionalities are limited, future iterations will incorporate user authentication features, allowing users to access personalized settings and account management options. This panel serves as a centralized hub for users to configure their app preferences and access essential information, contributing to a seamless and user-centric experience within the Citizen Engagement App.

Future continuation of the developed method is foreseen in the adaptation and configuration of additional information channels, e.g. wild animal and human negligence reporting are in the preparation and testing phase. Information channel about tick reporting is being developed to help users to protect from the potentially dangerous pathogens transferred by the ticks.

CONCLUSION

This paper describes the methods for citizen engagement (Mojir, 2024) specifically incorporated with the mobile application described in the work. The content of the paper focuses solely on the citizen engagement methodologies related with the development of CEA mobile application. The current state of related existing applications focuses on warning about the forest fires and a bit neglect the prevention from the forest fires. Except the Watch Duty and Fire Atlas, the applications focus more on data sharing than on gathering relevant data from mobile users. The content of the CEA distinguishes also with the variety of useful content (practical tips and guides) for situations which may happen during outdoor activities. We also believe that providing more functionality with adaptable content will attract more people for the real use of this application. CEA application is focused on sharing the information about wildfires and other hazardous situations within a community of subscribed users with specific role of the fire and rescue authorities which supervise the crowdsourced data and share validated information (fire hazard areas and escape routes) with citizens.

ACKNOWLEDGMENTS

This publication is the result of the following projects implementation: Integrated Technological and Information Platform for wildfire Management (SILVANUS) H2020 Green Deal Project No. 101037247; APVV No. APVV-22-0372 INFOTICK (Getting the right info on ticks); and Semantic Distributed Computing Continuum for Extreme Data Processing (2023-2025) VEGA 2/0131/23.

REFERENCES

Bailis, P., & Kingsbury, K. (2014). The network is reliable: An informal survey of real-world communications failures. Queue, 12(7), 20-32, https://dl.acm.org/doi/pdf/10.1145/2639988.2655736.

Balogh, Z.; Gatial, E.; Dolatabadi, S.H.; Dlugolinský, Štefan; Saltarella, M.; Scipioni, M.P.; Grunwald, D.; Przybyszewski, M.P.; Majlingova, A.; Brodrechtova, Y.; Mojir, K.Y.; Olson, N.; Levak, J.; Chandramouli, K., Communication Protocol for using Nontraditional Information Sources between First Responders and Citizens during Wildfires, Proceedings of the 20th International ISCRAM Conference (2023): Technologies for First Responders, 2514, pp. 152-165, http://dx.doi.org/10.59297/YIJK9098

Balogh, Z., Gatial, E., Hluchý, L., Poll Sourcing for Crisis Response. In A. Tapia, P. Antunes, V.A. Bañuls, K. Moore, & J. Porto (Eds.), ISCRAM 2016 Conference Proceedings, 13th International Conference on Information Systems for Crisis Response and Management: Emerging Topics, Rio de Janeiro, Brasil: Federal University of Rio de Janeiro

Bandura, A. (2001). Social cognitive theory of mass communication. Media Psychology, 3(3), 265–299, https://doi.org/10.1207/S1532785XMEP0303_03

Hano, M. C., Wei, L., Hubbell, B., & Rappold, A. G. (2020). Scaling Up: Citizen Science Engagement and Impacts Beyond the Individual. Citizen science: theory and practice, 5(1), 1-13, https://doi.org/10.5334/cstp.244

Hirschberger, P., Forests ablaze: Causes and effects of global forest fires, October 2016, available at: https://www.wwf.de/fileadmin/fm-wwf/Publikationen-PDF/WWF-Study-Forests-Ablaze.pdf

Jäckle, A., Burton, J., Couper, M., Lessof, C., Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: coverage and participation rates and biases, Survey Research Methods (2019), Survey Research Methods (2019), Vol. 13, No. 1, pp. 23-44, doi:10.18148/srm/2019.v13i1.7297

Kamilaris, A., Filippi, J.B., Padubidri, C., Koole, R., Karatsiolis, S., Examining the potential of mobile applications to assist people to escape wildfires in real-time, Fire Safety Journal, Volume 136, April 2023, https://doi.org/10.1016/j.firesaf.2023.103747

Messaoudi, C., Guessoum, Z., & BenRomdhane, L. (2022). Topic Extraction in Social Networks. COMPUTING AND INFORMATICS, 41(1), 56–77, https://doi.org/10.31577/cai_2022_1_56

Mojir, K.Y.; Maceviciute, E.; Olson, N.; Gatial, E.; Balogh, Z., Citizen Engagement in wildfire management: needs, challenges, methods and framework, Proceedings of the 20th International ISCRAM Conference (2023) : Track 10: Volunteers in Crisis Management/Emergency Response, 2564, pp. 761-772, http://dx.doi.org/10.59297/DGPE8556

Potter, W. (2012). Media effects. SAGE Publications, Chapter 3, https://dx.doi.org/10.4135/9781544308500

Romano, M., Onorati, T.,Aedo, I., Diaz, P., Designing Mobile Applications for Emergency Response: Citizens Acting as Human Sensors, Sensors 2016, https://doi.org/10.3390/s16030406

Sung, S. J., How Can We Use Mobile Apps for Disaster Communications in Taiwan: Problems and Possible Practice. 8th International Telecommunications Society (ITS) Asia-Pacific Regional Conference, 2011, pp. 1-15, available at: https://www.econstor.eu/handle/10419/52323

Teguh, R., Adji, F.F., Benius, Nur Aulia, M., Android mobile application for wildfire reporting and monitoring, Bulletin of Electrical Engineering and Informatics, Vol. 10, No. 6, December 2021, pp. 3412-3421, ISSN: 2302-9285

Wenz, A., Keusch, F., Increasing the Acceptance of Smartphone-Based Data Collection, Public Opinion Quarterly, Volume 87, Issue 2, Summer 2023, pp. 357-388, https://doi.org/10.1093/poq/nfad019

Towards a Taxonomy for Conversational Agents in Disaster Management

Johannes Göbel
Martin-Luther-University Halle-Wittenberg
johannes.goebel@wiwi.uni-halle.de

Hans Betke
Fraunhofer Institute FOKUS
hans.betke@fokus.fraunhofer.de

Stefan Sackmann
Martin-Luther-University Halle-Wittenberg
stefan.sackannn@wiwi.uni-halle.de

ABSTRACT

Conversational agents are a technology that is used today in many different ways, for example as chatbots or voice dialog systems. While they are mostly used for applications in the business sector, research is also focusing on their use in other areas, such as medicine or disaster management. Events in recent years, such as the global Covid pandemic and advances in the field of language learning models, have led to many new approaches. Taxonomies are a good way to provide researchers and practitioners with a good overview of this growing field of research by classifying new and existing approaches. In this paper we present the current results in a methodological approach to develop a taxonomy for the classification of conversational agent approaches in disaster management. We describe the data basis of a structured literature search, the implementation of the method and the current dimensions and characteristics of the emerging taxonomy.

Keywords

conversational agent, chatbot, taxonomy, classification, disaster management

INTRODUCTION

Software applications that use natural language processing to act as an interface between human users and information and communication systems (Dale, 2016) have been gaining in importance for years. These so-called conversational agents (CAs), also known as dialog systems, are already present in the everyday lives of many people, both in the form of voice assistants such as Alexa or Siri, as well as textually in the form of chatbots (Gnewuch et al., 2017). In 2018, there were already more than 300,000 chatbots in Facebook Messenger alone (Johnson, 2018), while the development and use of individual CAs by specialized providers such as Botpress or Voiceflow is becoming increasingly easy. Advances in the field of language learning models have led to the potential of this technology becoming increasingly utilizable (vom Brocke et al., 2018), with the result that there has recently been a veritable race between large IT companies such as Microsoft, Google and Meta to implement new applications in the field of CAs (Weise et al., 2023).

While many approaches of the past focused primarily on application areas in the fields of customer services or entertainment, new approaches are becoming increasingly diversified (Brandtzaeg & Følstad, 2017; Stieglitz et al., 2022). In the field of disaster and crisis management, CAs are also becoming an increasingly important technology that can be used for general information and communication tasks (e.g. Maniou & Veglis 2020), as well as for specific scenarios such as the Covid-19 pandemic (e.g. Rodswang et al. 2020). The constantly growing number of approaches in this field of application makes it increasingly difficult for scientists and practitioners to maintain an overview and distinguish key features.

To get a better overview of a research or application area and to classify existing approaches, taxonomies are a frequently used solution. Taxonomies are artifacts used in various scientific disciplines allow sequential classification of objects based on dimensions (Marradi, 1990). Several taxonomies have already been developed to classify CAs approaches from different perspectives. Janssen et al. (2020) focus on domain-specific chatbot applications, whereas Weber et al. (2021) present a taxonomy for CAs in education. Additional contributions provide conceptualizations of certain aspects of these systems like the taxonomy for social cues in CAs of Feine et al. (2019), as well as the exploration of chatbot relationship archetypes based on the time of interaction of Nißen et al. (2022).

However, to the best of our knowledge, there is no classification of CA approaches in the emerging area of disaster management. To address this research gap, in this paper we propose a taxonomy to classify approaches on conversational agents in disaster management. The results presented here are based on the methodological approach according to Nickerson et al. (2013) for the iterative development of taxonomies. By applying this method the following research question is to be specifically investigated in this research.

RQ: *Which dimensions and characteristics of a taxonomy describe current implementations of conversational agents in disaster management?*

However, as the research is still ongoing, the method has not yet been fully applied. The data basis for this taxonomy is provided by a structured literature analysis according to vom Brocke et al. (2009). For the complete research, we want to expand the database to include applied, non-scientifically documented applications and carry out a robust evaluation. Nevertheless, even in its current state of development, the taxonomy can support scientists in classifying existing and new approaches and offer practitioners possible criteria for selecting suitable solutions for their purposes.

The paper is structured as follows: in the second section, the current state of research on CAs is discussed. The methodological approach and its instantiation are presented in the third section before the taxonomy is described in the fourth section. Limitations and desiderata are discussed in the conclusion.

ON THE STATE OF CONVERSATIONAL AGENTS

Natural language processing as the basis technology for CAs has been a relevant topic in the field of information systems research for many decades. One of the earliest related computer programs was ELIZA, developed at MIT in the 1960s (Weizenbaum, 1966). It simulated psychotherapeutic dialogues by recognizing and responding to simple patterns in user input. Despite its lack of real intelligence, ELIZA impressed with its skillful pattern recognition and is considered an early example of chatbots. Text-based chatbots are a type of CAs alongside voice-based systems such as voice assistants (e.g. Siri, Alexa) (Gnewuch et al., 2017). Another category of CAs are so called embodied agents, which are visualized characters that often use facial expressions and gestures to provide a better user experience (Brave et al., 2005). Recent developments in this space are numerous and diverse, spanning various directions. One of these innovations is conversational dashboard systems, which allow users to interact with complex data visualizations using natural language (Ruoff et al., 2021). Another are conversational co-pilots which Microsoft, for example, is currently including in its product range (Dastin, 2023), as well as Coding Assistants (Ross et al., 2023) which automate tasks, advice, as well as collaborate with users.

An important theoretical approach in research on CAs is the social response theory (Gnewuch et al., 2017; Wambsganss et al., 2021). This states that people interact with computers as they do with other people, unconsciously applying social rules such as politeness, self-disclosure and trust, based on social cues such as conversation, interactivity and social roles. An important principle of this theory is that of reciprocity, which states that people respond to computers socially by attempting to "reciprocate" or "match" the computer's interaction or behavior (Nass & Moon, 2000; Reeves & Nass, 1996).This means that designers of CAs should consciously plan the conversational style of their applications, as the social characteristics of these systems have a significant influence on their use (Chattaraman et al., 2019).

Progress in the field of CAs is currently particularly dynamic. Systems such as OpenAI's Chat GPT are experiencing impressive growth. One million users registered for the service within five days of its launch - only Meta's "Threads" could reach so many users quicker (*Infographic*, 2023). They can also be easily integrated into existing systems such as websites, social media and messaging platforms like Facebook (Johnson, 2018). The latest developments are integrations of these systems into search engines, such as Bing Chat (*The Verge*, n.d.).

Social media has been an important control center for disaster response at least since the Central European flood 2013 (Albris, 2018). There, CAs have the potential to broadcast trustworthy information from civil protection authorities to the public (Stieglitz et al., 2022). They can also help overcome language barriers and provide reliable real-time information (Ahmady & Uchida, 2020). In addition, CAs have the potential to support the coordination

of spontaneous helpers during a crisis and reduce the cognitive load on individuals (Gerstmann et al., 2019; M.-H. Tsai et al., 2019). Although research into the use of CAs in disaster management is still in its infancy and there is a lack of established design approaches, new approaches are constantly emerging and current findings indicate that the use of CAs in disaster management provide added value (Stieglitz et al., 2022). The aim of the research presented here is to take a deeper look at the developments to date to enable a classification of existing and new approaches in this rapidly growing field of research.

METHODOLOGICAL APPROACH

Taxonomy Development Following Nickerson et al. (2013)

To build our taxonomy for CAs in disaster management on a sound basis, we use the established methodology of Nickerson et al. (2013). This method prescribes a multi-stage, cyclical procedure with defined ending conditions for deriving dimensions and characteristics for a topic-specific taxonomy. The instantiation of the method in the context of our research, illustrated in Figure 1, is described below.

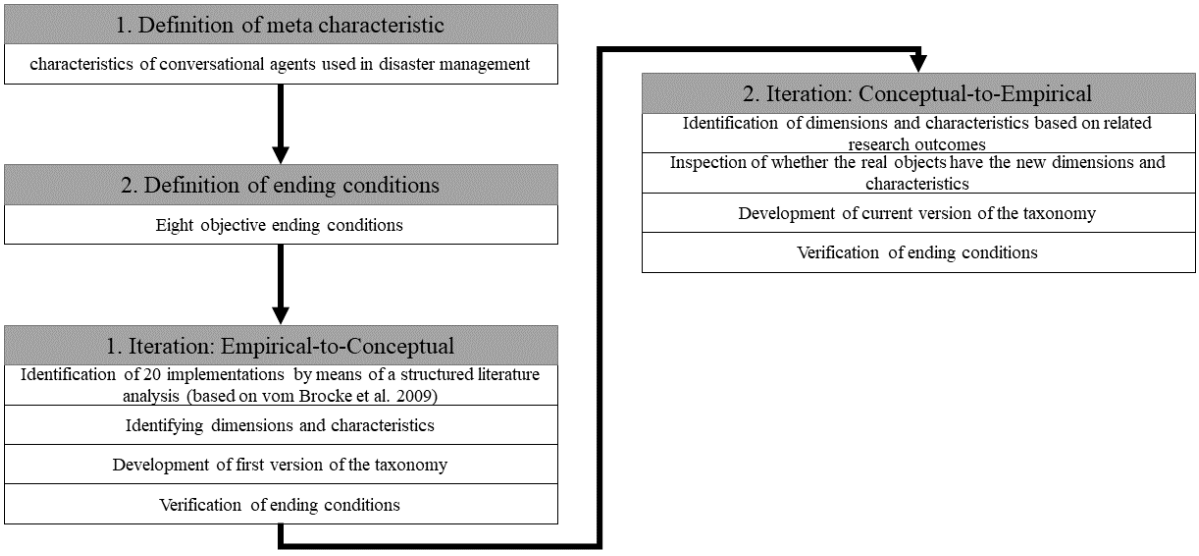


Figure 1. Procedure based on Nickerson et al. (2013)

The first step is the **definition of the meta-characteristic**. The aim here is to select a view of the objects to be examined (conversational agents in disaster management) from which all other characteristics can be derived (Nickerson et al., 2013). When defining the meta-characteristic, it is important to be clear about the objective and purpose of the taxonomy. In our case, the taxonomy is intended to be useful both for practitioners, who can use the taxonomy as a decision-making tool for the selection of suitable solutions, and for scientists from different disciplines who want to classify existing and new approaches and identify research gaps. The focus is on core features and application conditions, rather than specific technical or organizational aspects. The meta-characteristic was formulated in a correspondingly general manner: "characteristics of conversational agents used in disaster management."

The next step is to **define the ending conditions** of the iterative process for deriving new elements for the taxonomy. For this study the objective conditions (OEC) proposed by (Nickerson et al., 2013), which can be seen in Figure 2, are selected.

	Objective Ending Condition
1	All objects or a representative sample of objects have been examined
2	No object was merged with a similar object or split into multiple objects in the last iteration
3	At least one object is classified under every characteristics of every dimension
4	No new dimensions or characteristics were added in the last iteration
5	No dimensions or characteristics were merged or split in the last iteration
6	Every dimension is unique and not repeated (i.e., there is no dimension duplication)
7	Every characteristic is unique within its dimension (i.e., there is no characteristic duplication within a dimension)
8	Each cell (combination of characteristics) is unique and is not repeated (i.e., there is no cell duplication)

Figure 2. Objective ending conditions prescribed by (Nickeson et al., 2013)

In the following step the iterations for the development of characteristics and dimensions begin. These must be logical consequences of the meta-characteristic. A two-stage approach was chosen here, consisting of an initial "empirical-to-conceptual" step and a subsequent "conceptual-to-empirical" iteration. In the **first "empirical-to-conceptual" iteration** the required number of descriptions of CAs were created using a structured literature analysis, as recommended by Nickerson et al. (2013), to build the knowledge base. This is described in detail in the following section. A **second "conceptual-to-empirical" iteration** was then conducted to address dimensions related to important theoretical foundations, such as the Social Response Theory, which still need consideration. These are explored in categorizations of CAs in literature, as seen in the taxonomy presented by Janssen et al. (2020). Their taxonomy on domain-specific chatbots includes dimensions such as 'collaboration goal,' 'chatbot role,' 'relation duration,' and 'social cues' with their characteristics, which we adopt in this iteration (Janssen et al., 2020). These dimensions were used to examine the systems identified in the first cycle and assign them to the respective characteristics.

After the second iteration, the eight objective ending criteria were examined, we found that criterion OEC4 was not yet met, revealing that further iterations are necessary. As this paper is a presentation of research in progress, the procedure ends at this point. We are aware that further iterations are necessary, and an evaluation of the taxonomy will take place in future steps. The methodological instantiation described here forms the basis for the considerations on the preliminary taxonomy in the following main section.

Knowledge Base from Structured Literature Search

The first iteration of the taxonomy creation is based on a knowledge base consisting of a set of descriptions of CAs. Structured processing is recommended to set this up in a way that is comprehensible to third parties. We chose to build our knowledge base using a structured literature analyses following vom Brocke et al. (2009). The method includes a 5-step procedure: 1) definition of review scope, 2) conceptualization of topic, 3) literature search, 4) literature analysis, 5) research agenda (Brocke et al., 2009).

Starting with the **definition of review scope,** vom Brocke et al. propose a procedure based on the taxonomy of Cooper (1988) which was instantiated in Figure 3.

Characteristic	Categories			
focus	research outcome	research method	theories	applications
goal	integration	criticism	central issues	
organization	historic	conceptual	methodical	
perspective	neutral representation		espousal of position	
audience	specialized scholars	general scholars	practitioners/politicians	general public
coverage	exhaustive	exhaustive and selective	representative	central/pivotal

Figure 3. Instantiation of the Taxonomy for the "definition of review scope" according to Cooper (1988)

Based on this, the focus was placed on applications with the aim of addressing the central issues. The organization should be conceptual, whereby a neutral perspective should be adopted. Practitioners as well as specialized

researchers in the fields of disaster management and conversational systems were identified as the target group, who should be offered an exhaustive and selective coverage of the topic.

In the next phase, the **conceptualization of topic**, working definitions of the most important terms were created and an overview of the existing literature in the field of CAs in disaster management was compiled. The findings from this step were used to develop a better understanding and a relevant search string.

As part of the **literature search,** the scientific databases SpringerLink, ISCRAM, ScienceDirect, Wiley and Google Scholar were searched. Publications containing the search string in abstract, title or keywords were to be identified within the search. This is *(chatbot OR "conversational agent" OR "chat bot" OR "dialogue system") AND ("disaster" OR "crisis" OR "emergency")* and was formed from the core terms derived in the previous steps. In the Wiley and ScienceDirect databases, the search described above could be carried out via the website. For the other databases, additional steps had to be taken. The SpringerLink database had to be addressed programmatically using an Application Protocol Interface (API) to perform a full-text search. The title, abstract and keyword searches were then carried out using a Python script. The ISCRAM database did not provide the desired search functionalities either, but since it specializes in information systems in disaster control, a full-text search was also carried out on the site using Google Scholar. Finally, a title search was carried out using the search string in Google Scholar. This database also offered no possibility to search in abstract or keywords.

After the search process, in the **literature analysis,** the titles and abstracts of the results were checked according to previously defined exclusion criteria. These are (1) not in English, (2) not peer-reviewed, (3) no concrete description of a CA artifact and (4) the described artifact has not been implemented at least once. This process resulted in a total of 88 identified publications with 20 descriptions of specific CA implementations, which formed the knowledge base for the taxonomy's development. The entire literature search and the results of the individual analysis can be seen in Figure 4.

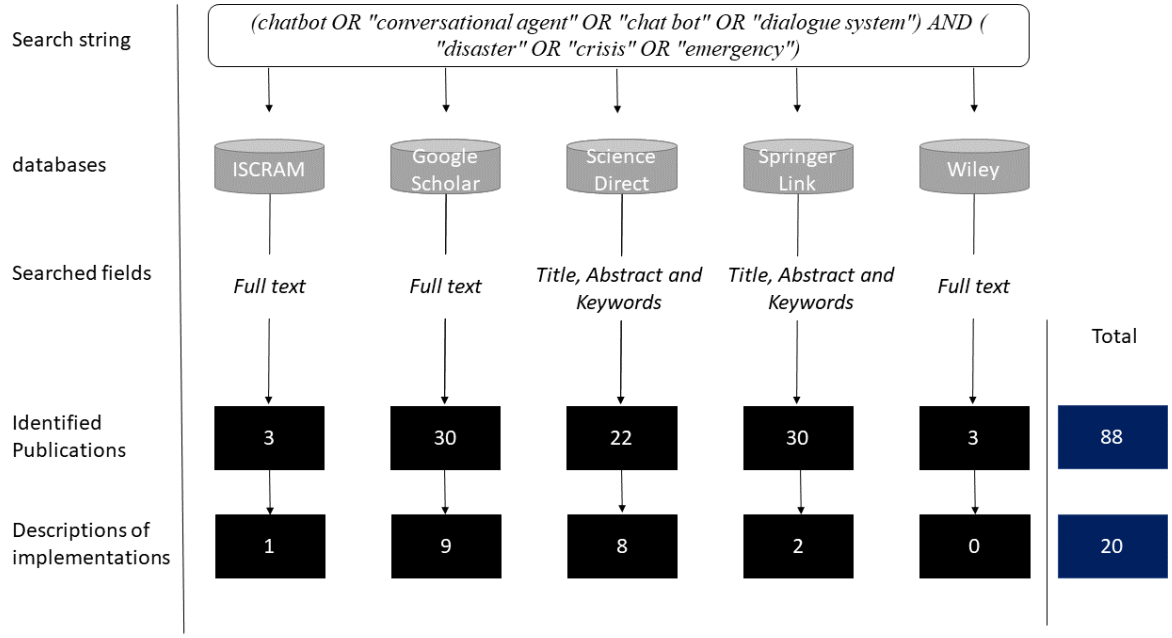


Figure 4. Literature Search Processing and Results

In this paper, the fifth step, formulation of a research agenda, is omitted as the focus on the methodology for taxonomy design renders it redundant.

A TAXONOMY FOR CONVERSATIONAL AGENTS IN DISASTER MANAGEMENT

This section describes the taxonomy in its current state of research. As part of the taxonomy development, 13 dimensions were identified. For reasons of clarity, these are divided into four categories and will be presented here. The developed taxonomy can be seen in Figure 5. Six of the included dimensions are exclusive, these are marked in the figure with an asterisk (*). The numbers in brackets after the characteristics indicate how many of the CA approaches from the knowledge base they apply to. Due to incomplete descriptions of the approaches, not all of them could be classified in every dimension.

Dimension		Characteristics			
Application Background	User*	Hospitals (2)	Disaster Management Authority (5)	Public Bodies (6)	Private Sector (2)
	Target Group*	General Population (12)	Affected (2)	Specific groups of the general population (4)	Disaster Management Officer (2)
	Type of Disaster*	Flood (4)	Pandemic (12)	Unspecific (4)	
Communication Goals and Language Processing	Intended Use	Coordination (6)	Information (13)	Education (1)	Decision Support (3)
	Conversational Goal*	Task-oriented (9)		Chatter (11)	
	Language Comprehension	Rule-based (7)		Natural (15)	
Interaction	Input Interaction	Text-based (16)	Voice-based (3)	Buttons (5)	Multimedia (2)
	Output Interaction	Text-based (16)	Voice-based (4)	Multimedia (6)	Facial Expression(1)
	Integration	Messenger Platform (10)	Native (8)	Smart Home Device (1)	
		Telephone (1)		Augmented Reality Device (1)	
	Human in the Loop	Forwarding to a Hotline (2)	Request to Adapt the Content (1)	No Described Functionality (17)	
Social Aspects	Role Model	Facilitator (9)	Peer (2)	Expert (9)	
	Relationship Length*	Short-term Relationship (12)		Long-term Relationship (8)	
	Social Cues*	Present (9)		Not Present (7)	

Figure 5. Taxonomy for Conversational Agents in Disaster Management

Category: Application Background

This category is intended to summarize all dimensions that deal with the primary actors and the overall situation in which the CA is used.

The **user** dimension describes the people and organizations that operate and administer the system. *Hospitals* are an example of this, but CAs are also used in the *private sector* (Jang & Lee, 2023; Sarbay et al., 2023). In contrast, the system designed for *disaster management authorities* by Tsai et al. attempts to provide a simple interface to retrieve up-to-date information on the current situation (M.-H. Tsai et al., 2019). A final characteristic is users in *public bodies*, such as government agencies or universities (Balderas et al., 2023).

The **target group** dimension describes which user group should be reached by the system. This can be the *general population*, whereas other systems try to address *disaster management officers* (Chan & Tsai, 2019; Nomura et al., 2020). The user group *Specific groups of the general population* can be illustrated by the "Emergency Reporter" system, which was created with the participation of indigenous groups to give them an opportunity to interact with the disaster management authorities (C.-H. Tsai et al., 2023). *Affected people* are also identified as a target group by systems (Syed et al., 2020).

The **disaster type** dimension describes the type of crisis in question, with a distinction being made between various types. A first example is the *flood*, but many CAs have also been identified that deal with *pandemics*, especially the Covid-19 pandemic (Sermet & Demir, 2018; Yoneoka et al., 2020). However, there are also systems that can be used *unspecifically* in several different types of crises. One example of this is DisBot, which uses knowledge accessed via social media to support citizens and first responders in disaster scenarios, improving community resilience and decision-making (Boné et al., 2020).

Category: Communication Goals and Language Processing

Strategic intentions of the system and the planned results of the interaction are relevant factors of CAs.

The **Intended Use** dimension refers to the purpose for which the CA was developed and which specific tasks or functions it should fulfill. The *Information* characteristic emphasizes the role of the CA as a source of information. The agent can for example provide real-time updates on hazardous situations, evacuation instructions and other relevant information to keep the public informed (Sermet & Demir, 2018). Used for *coordination*, a CA can ensure efficient communication and resource allocation during a disaster situation (Konstantoudakis et al., 2023). *Decision support* is another application in which the CA was developed to support decision-makers in disaster management with relevant data and semi-automated processes (Lai et al., 2020). Finally, a CA can be used in the field of *education* to prepare for disaster situations. By providing training materials, the agent helps to raise awareness and prepare for potential disasters (Syed et al., 2020).

The dimension of the **Conversational Goal** refers to the overarching goals or intentions that are pursued during the interaction with the user. A distinction is made here between CAs that act in a *task-oriented* manner and collaborate with the user (Tsai et al., 2023), and those that are *chatter*, as is the case with a simple question-answer scheme, for example (Goh et al., 2006; Konstantoudakis et al., 2023).

The **Language Comprehension** dimension of CAs describes how they process the user's utterances. *Natural language understanding* enables a CA to interpret the natural language of the user. This method enables versatile and human-like communication, as the agent can react to different expressions and formulations (Syed et al., 2020). In contrast, *rule-based language understanding* is based on predefined rules or scripts that the CAt uses to understand and respond to user utterances (C.-H. Tsai et al., 2023).

Category: Interaction

Interaction with a CA can be designed in very different ways. Dimensions in this category describe various aspects of the exchange of information between CAs and users.

The **Input Interaction** dimension describes how users can make inputs. Common input interactions take place via *text* or *voice* (Chan & Tsai, 2019; Nomura et al., 2020). An example of interaction with *buttons* is the chatbot presented by Maniou and Veglis for news dissemination in crises, which allows navigation using various button menus (Maniou & Veglis, 2020). Other systems can process *multimedia* input, for example a chatbot to collect information from volunteers and citizens, with the submission of images being one use case (Konstantoudakis et al., 2023).

The **Output Interaction** dimension specifies how output is transmitted to the user. Three types represent the output of *text*, *voice* or *multimedia* (Lai et al., 2020; Rodsawang et al., 2020; Yoneoka et al., 2020). A final type is the output of *facial expressions*, such as the intelligent embodied CA called AINI as a front-end interface for a crisis communication network portal (Goh et al., 2006).

The **Integration** dimension plays a crucial role in the embedding and deployment of CA systems, as it considers the different platforms and environments. These systems can be integrated in a variety of ways to ensure an optimal user experience. Examples of integration possibilities include implementation in *messenger platforms* to enable seamless interaction via messaging services (Konstantoudakis et al., 2023). Alternatively, a CA can be natively integrated into specific applications or operating systems to ensure deeper integration. Use via phone calls enables broader accessibility, while integration into smart home systems enables interaction in the context of the home environment (Lai et al., 2020; Sermet & Demir, 2018). In addition, integration into augmented reality platforms offers an immersive and enhanced level of interaction for users (Sermet & Demir, 2018).

The **Human in the Loop** dimension describes the possibility and mechanisms for human operators to intervene in the process. The *forwarding to a hotline* represents an identified mechanism that can be found in a voice Bot, for example. This partially automated the medical history processes for a hospital during the Covid-19 pandemic. As soon as it detected symptoms that indicated a disease, it forwarded the patient to a human operator (Lai et al., 2020). The *content customization prompt* was identified in an employee training application. The flow of conversation focuses on material created by a supervisor and can be interrupted if it is unclear. In such cases, a message is automatically sent to the supervisor, who is given the opportunity to revise the learning material (Syed et al., 2020). However, most approaches provide *no specific functionality*.

Category: Social Aspects

As described above, social response theory plays an important role in the design of CA. Dimensions of this category describe which general features CAs support in order to imitate human-like behavior.

The dimension **Role Model** classifies the existing systems based on the roles they take on towards the user. A CA in the role of a *facilitator* acts as a link between different parties to transmit information or facilitate communication (C.-H. Tsai et al., 2023). In the role of a *peer,* the CA interacts with the user on an equal footing without assuming a superior position. The interaction takes place in a cooperative and supportive manner, with the agent acting as a partner (Jiang et al., 2022). A conversational agent in the role of an *expert* imparts specialized knowledge and provides in-depth information on a specific topic (Goh et al., 2006)

The **Relationship Length** dimension concerns the duration and development of the interaction between the user and the CA, with different systems taking different approaches to the design of this relationship. An example of *short-term relationships* is a chatbot for medical history taking in a hospital presented by (Rodsawang et al., 2020). In contrast, Jiang et al. provide an example of a conversational agent that establishes a *long-term relationship.* They describe the use of the REPLIKA service during the Covid-19 pandemic which is intended to provide users with a supportive and empathetic conversation partner (Jiang et al., 2022).

Another dimension is the CA's ability to recognize, interpret and respond appropriately to **Social Cues**. These include facial expressions, gestures and tone of voice, which play a key role in creating an authentic and human-like interaction. An example of *present* social signals in this context is the output of facial expressions into a chatbot that can respond to user emotions (Goh et al., 2006). In contrast, these social signals are *not present* in other CAs (M.-H. Tsai et al., 2019).

CONCLUSION

In this paper, the research in progress and first results towards a taxonomy for conversational agents in disaster management was presented. The taxonomy was developed on the basis of an established methodology using an extensive, systematically compiled knowledge base. After two methodological iterations, a taxonomy consisting of 13 dimensions and 42 characteristics was created. The derived dimensions address both core features of the CAs and organizational aspects such as purpose and user groups, as well as domain-specific characteristics. This general orientation makes it possible to classify all domain-specific approaches and serve different purposes. The taxonomy can be used to help practitioners differentiate between suitable solutions for their purposes and to get an overview of what features they can or cannot expect from current IT systems. The examples given in the explanation of the dimensions can also provide inspiration for the use of CAs in disaster management. On the other hand, the taxonomy can also support scientists in the classification of research artifacts in the form of software demonstrators and application concepts. In addition, the taxonomy also represents an analysis of the current state of the art and shows which aspects of CAs for disaster management have so far been addressed to a greater or lesser extent by research. Although the analysis of the state of the art is not the focus of the paper, some trends can already be identified based on the frequency of the dimensions addressed. Most of the conversational agents examined serve to inform users and are generally designed for pandemics offering text-based interaction. The target group is usually the general population and subgroups of this population, with disaster managers being a secondary target. An interesting observation is that, in contrast to applications in the commercial sector only a few exceptions rely on human-in-the-loop functionalities. Even though research has shown that implementing human-in-the-loop functionalities in CAs can benefit users, task workers, and system performance (Reitmeier et al., 2020).

Since the research is still in progress not all steps in the underlying methodological approach of Nickerson et al (2013) have been executed yet. An important future step is the evaluation of the results. Nickerson et al. suggest having potential users evaluate the usefulness of the artifact. In the present scenario, disaster managers should be consulted to determine the extent to which the presented taxonomy aids them in choosing the appropriate CA for their specific use case. Similarly, the subjective termination criteria could not be considered and not all objective ending conditions could be met after the last iteration (Nickerson et al., 2013), so that further iterations will be necessary in future steps. The underlying knowledge base could be expanded for this purpose. With the selected search string, only CAs that were built specifically for disaster management applications can be identified. An extension of the search string could include other, domain-unspecific applications based on their potential use in disaster situations. The inclusion of solutions in practical use that have not yet been scientifically documented is also conceivable. In these ways, we want to further develop the taxonomy, which already offers decision-making and orientation aids for practitioners and scientists in its current state, into a valid and well-founded tool in the future.

AKNOWLEDGEMENT

This publication was produced in connection with the research project “KatHelfer-PRO”, funded by the German Ministry of Education and Research (grant number 13N16549).

REFERENCES

Ahmady, S. E., & Uchida, O. (2020). Telegram-based chatbot application for foreign people in Japan to share disaster-related information in real-time. *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, 177-181. https://ieeexplore.ieee.org/abstract/document/9118510/

Albris, K. (2018). The switchboard mechanism: How social media connected citizens during the 2013 floods in Dresden. *Journal of Contingencies and Crisis Management*, *26*(3), 350-357. https://doi.org/10.1111/1468-5973.12201

Balderas, A., García-Mena, R. F., Huerta, M., Mora, N., & Dodero, J. M. (2023). Chatbot for communicating with university students in emergency situation. *Heliyon*, *9*(9). https://www.cell.com/heliyon/pdf/S2405-8440(23)06725-7.pdf

Bing is now the default search for ChatGPT - The Verge. (n.d.). Retrieved January 8, 2024, from https://www.theverge.com/2023/5/23/23733189/chatgpt-bing-microsoft-default-search-openai-build

Boné, J., Ferreira, J. C., Ribeiro, R., & Cadete, G. (2020). Disbot: A Portuguese disaster support dynamic knowledge chatbot. *Applied Sciences*, *10*(24), 9082.

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. In I. Kompatsiaris, J. Cave, A. Satsiou, G. Carle, A. Passani, E. Kontopoulos, S. Diplaris, & D. McMillan (Eds.), *Internet Science* (Vol. 10673, pp. 377-392). Springer International Publishing. https://doi.org/10.1007/978-3-319-70284-1_30

Brave, S., Nass, C., & Hutchinson, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International journal of human-computer studies*, *62*(2), 161-178.

Chan, H.-Y., & Tsai, M.-H. (2019). Question-answering dialog system for emergency operations. *International Journal of Disaster Risk Reduction*, *41*, 101313.

Chattaraman, V., Kwon, W.-S., Gilbert, J. E., & Ross, K. (2019). Should AI-Based, conversational digital assistants employ social- or task-oriented interaction style? A task-competence and reciprocity perspective for older adults. *Computers in Human Behavior*, *90*, 315-330. https://doi.org/10.1016/j.chb.2018.08.048

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, *1*(1), 104. https://doi.org/10.1007/BF03177550

Dastin, Jeffrey: *Microsoft unveils AI office Copilot in fast-moving race with Google* In: *Reuters*, 16.03.2023. Retrieved 14. 02.2024.

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, *132*, 138-161.

Gerstmann, S., Betke, H., and Sackmann, S. (2019). Towards Automated Individual Communication for Coordination of Spontaneous Volunteers. In *Proceedings of the 16th Annual Global Conference on Information Systems for Crisis Response and Management (ISCRAM 2019), Spain*.

Gnewuch, U., Morana, S., & Maedche, A. (2017). *Towards Designing Cooperative and Social Conversational Agents for Customer Service*. Proceedings of the Thirty Eight International Conference of Information Systems (ICIS 2017).

Goh, O. S., Fung, C. C., Wong, K. W., & Depickere, A. (2006). An embodied conversational agent for intelligent web interaction on pandemic crisis communication. *2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops*, 397-400. https://ieeexplore.ieee.org/abstract/document/4053278/

Infographic: Threads Shoots Past One Million User Mark at Lightning Speed. (2023). Statista Daily Data. https://www.statista.com/chart/29174/time-to-one-million-users

Jang, H., & Lee, S. (2023). Introducing the Co-oriented Scansis (CoS) model: A case of chatbot, Lee-Luda. *Public Relations Review*, *49*(4), 102360. https://doi.org/10.1016/j.pubrev.2023.102360

Janssen, A., Passlick, J., Rodríguez Cardona, D., & Breitner, M. H. (2020). Virtual Assistance in Any Context. *Business & Information Systems Engineering*, *62*(3), 211-225. https://doi.org/10.1007/s12599-020-00644-1

Jiang, Q., Zhang, Y., & Pian, W. (2022). Chatbot as an emergency exist: Mediated empathy for resilience via human-AI interaction during the COVID-19 pandemic. *Information Processing & Management*, *59*(6), 103074.

Johnson, K. (2018). Facebook Messenger passes 300,000 bots. *VentureBeat*. https://venturebeat.com/ai/facebook-messenger-passes-300000-bots/

Konstantoudakis, K., Christaki, K., Sainidis, D., Babic, I., Kogias, D. G., Inglese, G., Bruno, L., Giunta, G., Patrikakis, C. Z., Balet, O., Rossi, C., Dimou, A., & Daras, P. (2023). Common Operational Picture and Interconnected Tools for Disaster Response: The FASTER Toolkit. In H. J. Scholl, E. E. Holdeman, & F. K. Boersma (Eds.), *Disaster Management and Information Technology: Professional Response and Recovery Management in the Age of Disasters* (pp. 83-110). Springer International Publishing. https://doi.org/10.1007/978-3-031-20939-0_6

Lai, L., Wittbold, K. A., Dadabhoy, F. Z., Sato, R., Landman, A. B., Schwamm, L. H., He, S., Patel, R., Wei, N., Zuccotti, G., Lennes, I. T., Medina, D., Sequist, T. D., Bomba, G., Keschner, Y. G., & Zhang, H. (Mark). (2020). Digital triage: Novel strategies for population health management in response to the COVID-19 pandemic. *Healthcare*, *8*(4), 100493. https://doi.org/10.1016/j.hjdsi.2020.100493

Maniou, T. A., & Veglis, A. (2020). Employing a chatbot for news dissemination during crisis: Design, implementation and evaluation. *Future Internet*, *12*(7), 109.

Marradi, A. (1990). Classification, typology, taxonomy. *Quality and Quantity*, *24*(2). https://doi.org/10.1007/BF00209548

Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, *56*(1), 81-103. https://doi.org/10.1111/0022-4537.00153

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*(3), 336-359. https://doi.org/10.1057/ejis.2012.26

Nißen, M., Selimi, D., Janssen, A., Cardona, D. R., Breitner, M. H., Kowatsch, T., & von Wangenheim, F. (2022). See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons. *Computers in Human Behavior*, *127*, 107043.

Nomura, S., Yoneoka, D., Shi, S., Tanoue, Y., Kawashima, T., Eguchi, A., Matsuura, K., Makiyama, K., Ejima, K., Taniguchi, T., Sakamoto, H., Kunishima, H., Gilmour, S., Nishiura, H., & Miyata, H. (2020). An assessment of self-reported COVID-19 related symptoms of 227,898 users of a social networking service in Japan: Has the regional risk changed after the declaration of the state of emergency? *The Lancet Regional Health - Western Pacific*, *1*, 100011. https://doi.org/10.1016/j.lanwpc.2020.100011

Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Center for the Study of Language and Information, Cambridge University Press.

Reitmaier, T., Robinson, S., Pearson, J., Kalarikalayil Raju, D., Jones, M. (2020). An Honest Conversation: Transparently Combining Machine and Human Speech Assistance in Public Spaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).

Rodsawang, C., Thongkliang, P., Intawong, T., Sonong, A., Thitiwatthana, Y., & Chottanapund, S. (2020). Designing a competent chatbot to counter the COVID-19 pandemic and empower risk communication in an emergency response system. *OSIR Journal*, *13*(2). http://www.osirjournal.net/index.php/osir/article/view/193

Ruoff, M., Gnewuch, U., Maedche, A., & Scheibehenne, B. (2023). Designing conversational dashboards for effective use in crisis response. *Journal of the Association for Information Systems*, *24*(6), 1500-1526.

Steven I. Ross, Fernando Martinez, Stephanie Houde, Michael Muller, and Justin D. Weisz. (2023). The Programmer's Assistant: Conversational Interaction with a Large Language Model for Software Development. In 28th International Conference on Intelligent User Interfaces (IUI '23), March 27–31, 2023, Sydney, NSW, Australia. ACM, New York, NY, USA, 24 pages. https://doi.org/10.1145/3581641.3584037

Sarbay, İ., Berikol, G. B., & Özturan, İ. U. (2023). Performance of emergency triage prediction of an open access natural language processing based chatbot application (ChatGPT): A preliminary, scenario-based cross-sectional study. *Turkish Journal of Emergency Medicine*, *23*(3), 156.

Sermet, Y., & Demir, I. (2018). An intelligent system on knowledge generation and communication about flooding. *Environmental Modeling & Software*, *108*, 51-60. https://doi.org/10.1016/j.envsoft.2018.06.003

Stieglitz, S., Hofeditz, L., Brünker, F., Ehnis, C., Mirbabaie, M., & Ross, B. (2022). Design principles for conversational agents to support Emergency Management Agencies. *International Journal of Information Management*, *63*, 102469. https://doi.org/10.1016/j.ijinfomgt.2021.102469

Syed, H. A., Schorch, M., & Pipek, V. (2020). Disaster learning aid: A chatbot centric approach for improved organizational disaster resilience. *Information Systems for Crisis Response and Management Conference; Virginia Tech: Blacksburg, VA, USA*, 448-457. https://kontikat.de/wp-

content/uploads/2021/07/Syed_Schorch_Pipek_2020_DisasterAid_ISCRAM2020.pdf

Tsai, C.-H., Rayi, P., Kadire, S., Wang, Y.-F., Krafka, S., Zendejas, E., & Chen, Y.-C. (2023). Co-Design Disaster Management Chatbot with Indigenous Communities. *20th International Conference on Information Systems for Crisis Response and Management (ISCRAM), 2023*, 1. https://digitalcommons.unomaha.edu/isqafacpub/126/

Tsai, M.-H., Chen, J. Y., & Kang, S.-C. (2019). Ask Diana: A keyword-based chatbot system for water-related disaster management. *Water*, *11*(2), 234.

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. *ECIS 2009 Proceedings*. 161. https://aisel.aisnet.org/ecis2009/161

vom Brocke, J., Maaß, W., Buxmann, P., Maedche, A., Leimeister, J. M., & Pecht, G. (2018). Future Work and Enterprise Systems. *Business & Information Systems Engineering*, *60*(4), 357-366. https://doi.org/10.1007/s12599-018-0544-2

Wambsganss, T., Höch, A., Zierau, N., & Söllne, M. (2021). Ethical Design of Conversational Agents: Towards Principles for a Value-Sensitive Design. *Wirtschaftsinformatik 2021 Proceedings*. https://aisel.aisnet.org/wi2021/ZMethods/Track01/2

Weber, F., Wambsganss, T., Rüttimann, D., & Söllner, M. (2021). Pedagogical agents for interactive learning: A taxonomy of conversational agents in education. In *Forty-Second International Conference on Information Systems. Austin, Texas* (pp. 1-17).

Weise, K., Metz, C., Grant, N., & Isaac, M. (2023). Inside the A.I. Arms Race That Changed Silicon Valley Forever. *The New York Times.* https://www.nytimes.com/2023/12/05/technology/ai-chatgpt-google-meta.html

Weizenbaum, J. (1966). ELIZA-a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36-45. https://doi.org/10.1145/365153.365168

Yoneoka, D., Tanoue, Y., Kawashima, T., Nomura, S., Shi, S., Eguchi, A., Ejima, K., Taniguchi, T., Sakamoto, H., Kunishima, H., Gilmour, S., Nishiura, H., & Miyata, H. (2020). Large-scale epidemiological monitoring of the COVID-19 epidemic in Tokyo. *The Lancet Regional Health - Western Pacific*, *3*, 100016. https://doi.org/10.1016/j.lanwpc.2020.100016

Tweeting Through the Flood: Application of BERT Topic Modeling for a Comparative Flood Communication Analysis

Christin Salley
Michigan Institute for Data Science
cjsalley@umich.edu

Nathan Fox
Michigan Institute for Data Science
foxnat@umich.edu

Alyssa Schubert
Michigan Institute for Data Science
alschub@umich.edu

ABSTRACT

Floods are prevalent disasters in the United States, posing escalating risks due to climate change-induced factors like rising sea levels and erratic rainfall patterns. Despite governmental efforts, flood risk communication rema ins inadequate, hindering preparedness and response capacity. While governmental agencies predominantly employ traditional, one-sided information dissemination approaches, social media platforms pose as crucial early warning indicators and sources of real-time information. This study conducts the first part of our intended comparative analysis of social media messages between governmental agencies and communities. We perfo rm a case study on a flooding event in Michigan from May 17-20, 2020. Utilizing advanced topic modeling, we examine Twitter/X message content and sentiment from community-based posts. Insights aim to inform more effective flood communication strategies, bridging the gap between official information and community needs during events. Future work will compare social media messages' content and sentiment from governmental agencies relative to those from the community.

Keywords

Crisis informatics, flooding, natural language processing (NLP), social media, Twitter/X

INTRODUCTION AND BACKGROUND

Floods are one of the most common weather-related natural disasters in the United States (NOAA National Severe Storms Laboratory, 2024). Indeed, seventy-five percent of all Presidential Disaster Declarations in the United States can be attributed to flooding (National Weather Service, 2018). As climate change continues to contribute to rising sea levels and unstable rainfalls, the risk of floods is increasing (Tabari, 2020). In fact, the U.S. National Flood Insurance Program experienced a 660% increase in dollars paid out for flood insurance claims from 2000 to 2020 compared to the claims from 1980 to 2000 (Federal Emergency Management Agency, 2021). Beyond flood damage to physical systems, floods have been documented to substantially impact communities' physical, mental, and financial health, causing long-term changes to quality of life that can persist for years after the flood event (Van

CoRe Paper – Risk Communication and Community Engagement
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Ootegem & Verhofstadt, 2016). For example, studies have shown increased instances in anxiety, depression, and post-traumatic stress disorder in people who have experienced a flood compared to a control group (Van Ootegem & Verhofstadt, 2016). Further, those affected by floods have been reported to experience long- or short-term injuries, respiratory and gastrointestinal illness due to exposure to pathogens via contaminated drinking water, unsafe conditions for food preparation, or lack of access to hygienic practices, and, in some cases, carbon monoxide poisoning (Burger & Gochfeld, 2014; Mulder et al., 2019; Sampson et al., 2019; Van Ootegem & Verhofstadt, 2016; Waite et al., 2014). Moreover, these effects are not experienced independently of one another; a systematic review of factors related to one's vulnerability before, during, and after a flood event determined that the effects of floods can be compounded by other risk factors, such as the frequency and severity of floods experienced or health, socioeconomic, and educational status (Lowe et al., 2013; Sampson et al., 2019). In particular, urban flooding has been shown to occur disproportionately due to racial and spatial segregation (Hughes et al., 2021). Midwest states, including cities on the Great Lakes, are particularly prone to urban flooding, which tends to occur in cities with high percentages of impervious surfaces and outdated or aged infrastructure, such as combined stormwater/sewer systems (National Academies of Sciences, 2019).

The magnitude and widespread effects of floods, as well as their increasing risk of occurrence, has motivated a number of studies examining the information available to those experiencing a flood before, during, and after the flood event. The primary means through which communities receive information about a flood is via flood risk messages from a governmental agency, such as public health, weather service, or emergency management agencies on a federal, state, or local level. However, studies investigating the effectiveness of flood risk messages from these sources have found that they are inadequate for improving individual self-efficacy, adaptive capacity, and flood preparedness (Forsyth et al., 2023; Haer et al., 2016; Scott & Errett, 2018). For example, a study examining the information available on social media via government during the 2016 Louisiana floods found that, while most messages had information about recovery resources and updates about the status of the disaster and response, few contained 'actionable requests,' or requests that were substantial and relevant enough to those experiencing the flood that they increased self-efficacy (Cooper et al., 2022; Mostafiz et al., 2022; Scott & Errett, 2018). Further, while some agencies release 'Get Ready' checklists as a tool for preparedness before a flood, they do not adequately cover the long-term impacts on mental health and overall well-being that may be experienced (Forsyth et al., 2023). Current flood risk communication practices are still primarily based on a knowledge- or information-deficit model which perpetuates one-sided communication from a governmental agency to the community (Maidl & Buchecker, 2015). These models are based on the assumption that the public does not understand the risk and that simply providing more information will result in 'better' decision-making (Maidl & Buchecker, 2015; O'Sullivan et al., 2012).

An increasingly recognized additional source of flood risk information are the messages, communicated via social media, by the community who is affected by the flood. Indeed, in an effort to improve flood mitigation and management practices, several studies have demonstrated the effectiveness of community-based social media messages as early warning indicators of floods (Carlos Villagrán De León et al., 2013; Cools et al., 2016). For example, a study analyzing the content of 60,000 tweets over a 5-day flood period found evidence of early and local detection of multiple-flood events (Shoyama et al., 2021). A method for identifying flood events globally using social media messages to build a real-time and historic database has been created (de Bruijn et al., 2019). A recent study developed a method for real-time sentiment analysis of social media messages to provide up-to-date information on community needs during a flood (Bryan-Smith et al., 2023); another designed a multilingual multimodal neural network to classify tweets containing information related to a flood with high precision ((de Bruijn et al., 2020). While these studies use a variety of approaches to analyze the content and sentiment of social media messages, machine learning methods such as natural language processing (NLP) can considerably speed up the process. Topic modeling is an NLP technique that generates structured text data from unstructured text data inputs, and in doing so, allows for the discovery of latent topics in the text. Analyzing social media messages for content and sentiment using NLP or topic modeling has been well-established in the literature. In the disaster space, for example, topic modeling has been used to identify the sentiments and thoughts of populations during flood events (Choirul Rahmadan et al., 2020), hurricanes (Mihunov et al., 2022; Zhou et al., 2023), and earthquakes (Vo & Collier, 2013).

Few studies have completed a combined analysis of the social media messages from governmental agencies and community networks during a flood event. However, given the observed gaps in flood risk communication from governmental agencies, there is an opportunity to learn about the community's needs and concerns from the content

and sentiment of community-based social media messages in comparison to government-based social media messages during flood events. A better understanding of these needs can be used to build more effective communication practices and reconcile the gap between what is provided versus what is needed by communities before, during, and after a flood. To that end, this preliminary study conducts a comparative analysis of the content and sentiment of tweets from governmental agencies versus community residents for the state of Michigan during a flood event from May 17 - 20, 2020 using topic modeling. We aim to answer the following research questions: (1) What insights can be gained from a comparative study of these social media messages in the context and sentiments of flood response and communication strategies?, and (2) Do the frequency of specific topics, message content, and sentiments of social media messages during the flood event differ between governmental (i.e., state and federal) agencies and community residents?

RESEARCH DESIGN AND METHODS

Case Study

To investigate our research questions thoroughly, we performed a case study analyzing a flooding event that occurred in Michigan during the year 2020. Midland County, Michigan is located in close proximity to Saginaw Bay, which connects to the Great Lake Huron. On May 17, 2020, record rainfall in Midland County began due to a low pressure system and frontal boundary from the Great Lakes (National Weather Service, 2020). Consequently, the Saginaw and Tittabawasee Rivers flooded. Heavy rains, totaling 5 to 8 inches, in the Tri-Cities region (i.e., Bay County, Midland County, and Saginaw County) resulted in catastrophic failures of the Edenville and Sanford dams in Midland County. This prompted the issuance of Flash Flood Emergencies, signaling life-threatening flooding and the evacuation of 10,000 Midland residents. Significant flooding also occurred along several lake shores, affecting several other counties in southeast Michigan. The rainfall fell within thirty-six hours and the event was considered over on May 20, 2020. This flood event was the worst that had occurred in the region since 1986, prompting discussion by those throughout the state on social media.

Data Collection and Filtering

Twitter data were collected for the state of Michigan from between May 15, 2020 at 20:00:09 and May 23, 2020 at 19:59:57 using a Twitter/X streaming API built in the Network Dynamics Lab (Wang & Taylor, 2015). In the realm of disaster studies, flooding is a prime example of what is known as a short-notice disaster (Chiu & Zheng, 2006). This means agencies typically have only 24 to 72 hours to send out important risk communication messages, such as evacuation notices (Wolshon et al., 2001). Furthermore, upon our initial analysis, we observed a consistent decline in the frequency of tweets concerning our case study within three days following the conclusion of the event. Therefore, the date and time range of May 15, 2020 to May 23, 2020 were chosen to encompass the average timeframe for a short-notice disaster (i.e., 48 hours prior) prior to the beginning of the flood event and three days after the incident. To compare the content and sentiment of tweets between governmental agencies and the public, the Twitter data were filtered for twelve specific state and federal agency handles (Table 1). This list was determined by including the social media platforms for relevant Executive Branch departments in the state (State of Michigan, 2024), as well as common federal agencies employed for disastrous events such as floods. Tweets from these handles were treated as the messages from governmental agencies and all other remaining tweets from the data collection were considered to be from the public or community based.

Table 1. List of state and federal governmental agencies for comparative analysis.

Agency Name	Twitter Handle
Environmental Protection Agency (EPA)	@EPA
Federal Emergency Management Agency (FEMA)	@fema
Michigan Department of Environment, Great Lakes, and Energy (EGLE)	@MichiganEGLE
Michigan Department of Health and Human Services	@MichiganHHS
Michigan Department of Natural Resources	@MichiganDNR
Michigan Department of Transportation (MDOT)	@MichiganDOT
Michigan State Police (MSP)	@MichStatePolice
National Weather Service (NWS)	@NWS
National Weather Service (NWS)	@NWSFlashFlood
National Weather Service (NWS)	@NWSSevereTstorm
State Emergency Management Agency (Michigan State Police, Emergency Management and Homeland Security Division)	@MichEMHS
State of Michigan government agencies	@migovernment

Topic Modeling

Next, to analyze the Twitter/X dataset collected, BERT (Bidirectional Encoder Representations from Transformers), an advanced NLP model introduced by Google in 2018 (Devlin et al., 2018), was used. BERT utilizes bidirectional encoding to generate contextualized word embeddings, revolutionizing various NLP tasks. For this project, we specifically employed BERTopics, an extension of BERT, that leverages its contextualized embeddings to perform topic modeling by aggregating document-level representations derived from BERT embeddings and subsequently clustering similar documents based on their semantic content. This method enables fine-grained topic differentiation and accurate thematic analysis, and has been used successfully to assess Twitter/X datasets for a range of environmental topics (Kastrati et al., 2023; Pampapura Madali et al., 2022).

We carried out hyperparameter tuning to optimize the performance of our topic modeling approach. During this process, we tuned the parameters of three key models: UMAP (Uniform Manifold Approximation and Projection) for dimensionality reduction, HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) for clustering, and BERTopic for topic modeling. Specifically, we adjusted parameters such as the number of neighbors and components for UMAP, the minimum cluster size and minimum samples for HDBSCAN, and minimum topic size for BERTopic. For a comprehensive list of the tuned parameters, please refer to the Supplementary Information at the end of the paper (SI 1).

To understand the meaning of each topic, we manually interpreted them and then labeled them based on their keywords to best capture their core meaning. The topic modeling tool assigns a c-TF-IDF (class-based term frequency-inverse document frequency) score to each word in the topic, indicating its relevance to the topic. These scores are ranked by importance, with the highest-scoring words most closely mirroring the central theme of the topic. While more topics can contain up to ten words, those with the highest c-TF-IDF score better reflect the overall meaning of the topic.

ANALYSIS

Results: Entire Dataset

We extracted 124,362 tweets from May 15 to May 23. The topic modeling clustered the tweets into 133 topics. The four most tweeted topics from the entire dataset were about "Food" (Topic 0, 4,583 tweets), "Thankfulness" (Topic 1, 3,411), "Jobs" (Topic 2, 2,537), and "Music" (Topic 3, 2,402) (Figure 1).



Figure 1. Top most tweeted topics from May 15 to May 23 in Michigan.

The topic most related to flooding events in Midland was Topic 15, which had 892 tweets in the cluster. The five most important words in Topic 15 were "Midland", "Dam", "Flooding", "Water", and "River" (Figure 2). Figure 2 also demonstrates the decline in score c-TF-IDF for each additional word included in the topic. The frequency of tweets categorized as belonging to Topic 15 varied over the search period. To start, there were virtually no tweets about the flood from between May 16 and May 19, followed by a rapid increase and peak posting on May 20, and then a decline between May 20 and May 23; however, during this period, the Tweet frequency remained higher than before the event (Figure 3).

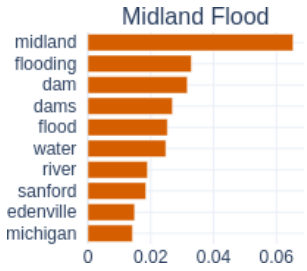


Figure 2. The top ten words associated with Topic 15, "Midland Flood".

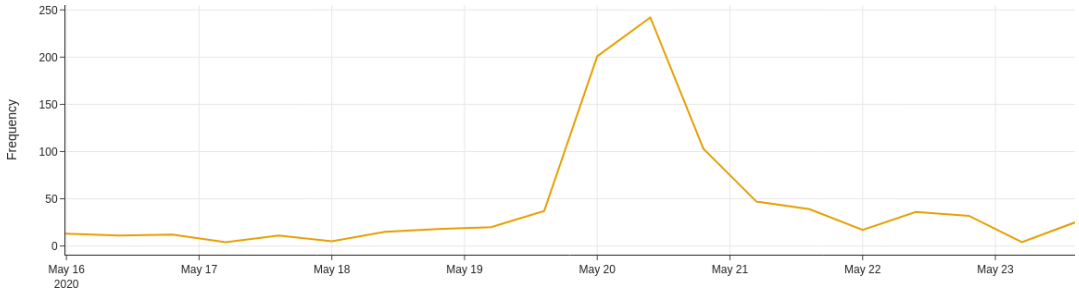


Figure 3. Frequency of Tweets associated with Topic 15 from May 16 to May 23, 2020.

The sentiment analysis results showed that for Topic 15 the most conveyed emotions were fear (347 tweets) and sadness (176) (Figure 4).

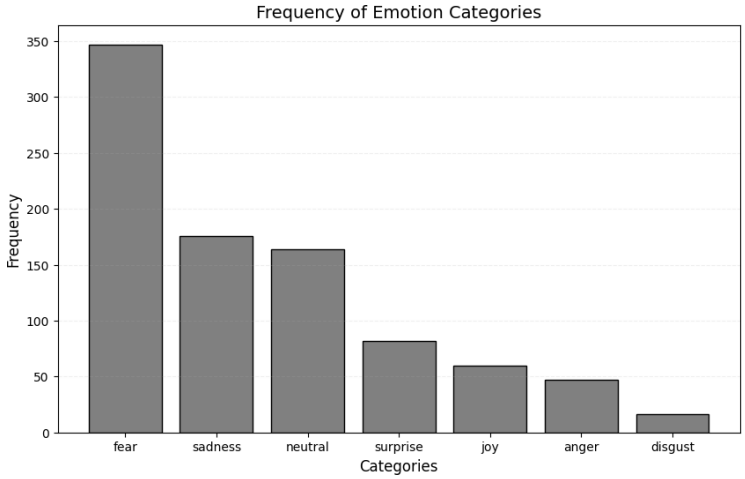


Figure 4. Sentiment Analysis of Topic 15 Tweets.

A hotspot analysis indicates that there is geographical variation in the location of tweets regarding the flooding events. Unsurprisingly, the location with the highest density of tweets was in and around Midland (Figure 5). There were also a large number of tweets occurring in other major Michigan cities such as Detroit and Lansing.

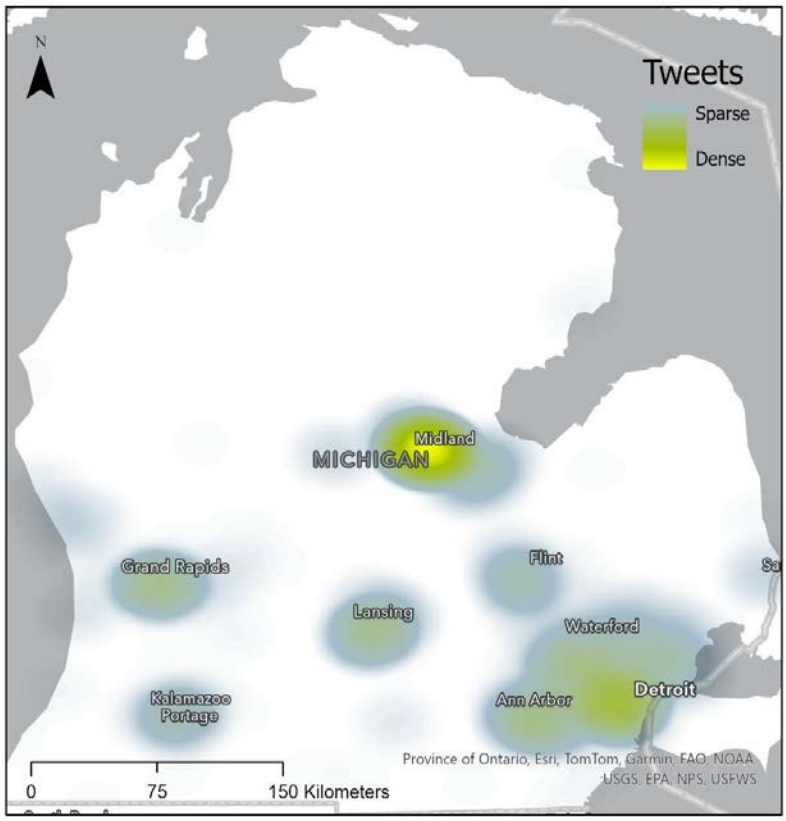


Figure. 5 Heatmap of tweets in the state of Michigan concerning the May 2020 flooding incident.

Results: Initial Comparison of Community vs. Government Tweets

Upon analyzing the dataset and dividing it into community and government tweets concerning the flooding event, we encountered a disproportionately small sample size of government tweets (13 out of 892 tweets; 1.4%) compared to the extensive volume generated by community members. All agency tweets were by either NWSDetoit (6 tweets) or NWSGaylord (5 tweets). Of these tweets, only three were directly related to Midland County, and a further three related to the associated flooding occuring in the village of Sanford, while all other tweets focused on a Gladwin County flooding event that occurred further north. This discovery perpetuates the ongoing notion of inadequate communication from key stakeholders at the local, state, and federal levels and the communities they serve, particularly on platforms as influential and interactive as social media, and identifies an opportunity to reach a broader audience through utilizing social media as a communication medium.

As our research remains ongoing, we intend to conduct another case study focusing on a different flooding event. Our aim is to either achieve a more robust sample size for both groups or pivot the study's focus to further comprehensively analyze the entire dataset, thereby uncovering novel insights in the forthcoming weeks. We anticipate presenting our findings at this conference upon acceptance and final submission of this work in progress paper.

DISCUSSION AND CONCLUSIONS

The particular insights gained from our study in the context and sentiments of flood response and communication strategies was the discovery that tweets surrounding the flood predominantly conveyed fear. Additionally, the frequency of tweets for Topic 15 (i.e., the topic most relevant to the incident itself) began rising mid-day on May 19, rapidly increasing on the 20th and peaking mid-day on the 21st. Overall, the highest volume of tweets for Topic 15 occurred between May 19 and May 22 with discussion continuing, to a lesser extent, on May 23. Rainfall began on May 17 and ended approximately thirty-six hours later. The temporal trends of tweet volume starting on May 19 shows, in near real-time, the community discussion and reactions to the impacts of the rainfall, including catastrophic flooding and dam breakages. Such insights are invaluable for stakeholders involved in risk communication, aiding in the development of initiatives and strategies to effectively disseminate information that promotes community safety during such disasters. This includes understanding not only the content to convey but also the timing when people are most likely to engage with and absorb such messages amidst the peak of discussion about the event.

Although this study effectively utilized the BERT model to identify communication inefficiencies, certain limitations must be acknowledged. Firstly, as is common in social media research, there exists variability in the collected tweets, including differences in the frequency of tweets and how engaged users are. Notably, the volume of community-scale tweets significantly outweighed those from governmental sources. It is possible that governmental agencies were communicating about the flood on another social media platform other than Twitter. Given the volume of community Tweets during the flood event, the lack of tweets from governmental sources indicate that there may be a missed opportunity to communicate important information about the flood via this platform. Consequently, the lack of a balanced sample size may restrict the generalizability of the conclusions to other flooding events. Additionally, the filtering process (i.e., the selected state and government entities) implemented during data collection may have influenced the results obtained, potentially impacting the comprehensiveness of the findings. For example, we have not captured any tweets that may have originated from national agencies because we filtered for tweets from the state of Michigan only (i.e., EPA and FEMA).

In future research, it is imperative to explore additional social media networks for government-based tweets. For instance, investigative reporters or news journalists that work on the local, state, and federal level and their networks can be included in future works to improve the dataset for flood risk communication from government stakeholders. Additionally, expanding the analysis to encompass tweets from local agencies can be advantageous. We initially concentrated on state and federal entities, and finding more local agencies would provide a more comprehensive understanding of flood-related communication patterns and potentially increase our dataset. To this end, future investigations should delve deeper into other forms of classification in the content of flood-related tweets, following the approach outlined by (Scott & Errett, 2018). For instance, examining tweets within specific topics, such as Topic 15, could shed light on the discussion of resource requirements, incidents of downed power lines, and other relevant themes. Lastly, previous studies have found different predominant emotion from tweets surrounding other natural hazard events such anger associated with hurricanes (Berbère et al., 2023). This disparity may show that different

natural hazard event types may evoke different emotion responses, or that public emotion may change throughout the timing of the event. As these studies only looked at the dominance of emotion and not temporal patterns, future analysis of such data should map the temporal distribution of emotion to best understand public perceptions. For example, fearful tweets may occur primarily before the event, whilst angry tweets may occur more frequently in the aftermath.

In conclusion, our work highlights how pressing challenges in flood risk communication persist, emphasizing the need for innovative approaches. While a more comparative analysis is forthcoming, the lack of government-based tweets generated again still show significant inefficiencies from these stakeholders to communicate with citizens. By analyzing social media messages, illustrated during a flooding event in Michigan, the study emphasizes the potential of social media platforms in enhancing real-time information dissemination and community engagement. These findings call for a paradigm shift towards more collaborative and responsive flood communication strategies to better address community needs and bridge the gap between official agencies and the public during flood events.

ACKNOWLEDGEMENTS

This project is supported by the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Futures program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors. We would also like to extend a special thank you to the Network Dynamics Lab at the Georgia Institute of Technology, for providing their archived Twitter/X postings for us to utilize in this research.

REFERENCES

Berbère, R., Elkefi, S., Layeb, S. B., & Tounsi, A. (2023). Exploring Cognitive Sustainability Concerns in Public Responses to Extreme Weather Events: An NLP Analysis of Twitter Data. *Cognitive Sustainability, 2(4)*.

Bryan-Smith, L., Godsall, J., George, F., Egode, K., Dethlefs, N., & Parsons, D. (2023). Real-time social media sentiment analysis for rapid impact assessment of floods. *Computers and Geosciences*, *178*. https://doi.org/10.1016/j.cageo.2023.105405

Burger, J., & Gochfeld, M. (2014). Health concerns and perceptions of central and coastal New Jersey residents in the 100days following Superstorm Sandy. *Science of the Total Environment*, *481*(1), 611–618. https://doi.org/10.1016/j.scitotenv.2014.02.048

Carlos Villagrán De León, J., Pruessner, I., & Breedlove, H. (2013). *Alert and Warning Frameworks in the Context of Early Warning Systems: A Comparative Review*.

Chiu, Y. C., & Zheng, H. (2007). Real-time mobilization decisions for multi-priority emergency response resources and evacuation groups: model formulation and solution. *Transportation Research Part E: Logistics and Transportation Review*, *43*(6), 710-736.

Choirul Rahmadan, M., Nizar Hidayanto, A., Swadani Ekasari, D., Purwandari, B., & Theresiawati. (2020). Sentiment Analysis and Topic Modelling Using the LDA Method related to the Flood Disaster in Jakarta on Twitter. *Proceedings - 2nd International Conference on Informatics, Multimedia, Cyber, and Information System, ICIMCIS 2020*, 126–130. https://doi.org/10.1109/ICIMCIS51567.2020.9354320

Cools, J., Innocenti, D., & O'Brien, S. (2016). Lessons from flood early warning systems. *Environmental Science & Policy*, *58*, 117–122. https://doi.org/10.1016/j.envsci.2016.01.006

Cooper, C. M., Sharma, S., Nicholas, R. E., & Keller, K. (2022). Toward More Actionable Flood-Risk Information. In *Earth's Future* (Vol. 10, Issue 11). John Wiley and Sons Inc. https://doi.org/10.1029/2022EF003093

de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J., & Aerts, J. C. J. H. (2019). A global database of historic and real-time flood events based on social media. *Scientific Data*, *6*(1). https://doi.org/10.1038/s41597-019-0326-9

de Bruijn, J. A., de Moel, H., Weerts, A. H., de Ruiter, M. C., Basar, E., Eilander, D., & Aerts, J. C. J. H. (2020). Improving the classification of flood tweets with contextual hydrological information in a multimodal neural network. *Computers and Geosciences*, *140*. https://doi.org/10.1016/j.cageo.2020.104485

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. http://arxiv.org/abs/1810.04805

Federal Emergency Management Agency. (2021). *Climate Change Increases Risk of Flooding*. https://www.floodsmart.gov/climate-change#:~:text=The%20risk%20of%20flooding%20is,and%20rising%20risk%20of%20flooding.

Forsyth, W., Roberts, T., & Brewer, G. (2023). Conceptualising risk communication barriers to household flood preparedness. *Urban Governance*, *3*(2), 116–129. https://doi.org/10.1016/j.ugj.2023.02.001

Haer, T., Botzen, W. J. W., & Aerts, J. C. J. H. (2016). The effectiveness of flood risk communication strategies and the influence of social networks-Insights from an agent-based model. *Environmental Science and Policy*, *60*, 44–52. https://doi.org/10.1016/j.envsci.2016.03.006

Hughes, S., Schwarz, K., Lemee, G. L., Dobie, S., Lane, M., Gonzalez, A., Turner, K., Norton, R., Baja, K., Howrani-Heeres, J., & Stults, M. (2021). *Centering Racial Justice in Urban Flood Adaptation: Planning and Evaluation Tools for Great Lakes Cities*.

Kastrati, Z., Imran, A. S., Daudpota, S. M., Memon, M. A., & Kastrati, M. (2023). Soaring Energy Prices: Understanding Public Engagement on Twitter Using Sentiment Analysis and Topic Modeling with Transformers. *IEEE Access*, *11*, 26541–26553. https://doi.org/10.1109/ACCESS.2023.3257283

Lowe, D., Ebi, K. L., & Forsberg, B. (2013). Factors increasing vulnerability to health effects before, during and after floods. *International Journal of Environmental Research and Public Health*, *10*(12), 7015–7067. https://doi.org/10.3390/ijerph10127015

Maidl, E., & Buchecker, M. (2015). Raising risk preparedness by flood risk communication. *Natural Hazards and Earth System Sciences*, *15*(7), 1577–1595. https://doi.org/10.5194/nhess-15-1577-2015

Mihunov, V. V., Jafari, N. H., Wang, K., Lam, N. S. N., & Govender, D. (2022). Disaster Impacts Surveillance from Social Media with Topic Modeling and Feature Extraction: Case of Hurricane Harvey. *International Journal of Disaster Risk Science*, *13*(5), 729–742. https://doi.org/10.1007/s13753-022-00442-1

Mostafiz, R. Bin, Rohli, R. V., Friedland, C. J., & Lee, Y. C. (2022). Actionable Information in Flood Risk Communications and the Potential for New Web-Based Tools for Long-Term Planning for Individuals and Community. *Frontiers in Earth Science*, *10*(February). https://doi.org/10.3389/feart.2022.840250

Mulder, A. C., Pijnacker, R., De Man, H., Van De Kassteele, J., Van Pelt, W., Mughini-Gras, L., & Franz, E. (2019). 'Sickenin'' in the rain" - Increased risk of gastrointestinal and respiratory infections after urban pluvial flooding in a population-based cross-sectional study in the Netherlands'. *BMC Infectious Diseases*, *19*(1). https://doi.org/10.1186/s12879-019-3984-5

National Academies of Sciences, E. and M. (2019). *Framing the Challenge of Urban Flooding in the United States*. The National Academies Press. https://doi.org/https://doi.org/10.17226/25381

National Weather Service. (2018). *Flood Related Hazards*. https://www.weather.gov/safety/flood-hazards

National Weather Service. (2020). *Historic Flooding May 17-20, 2020*. https://www.weather.gov/dtx/HistoricFlooding-May-17-20-2020

NOAA National Severe Storms Laboratory. (2024). *Severe Weather 101*. https://www.nssl.noaa.gov/education/svrwx101/floods/

O'Sullivan, J. J., Bradford, R. A., Bonaiuto, M., De Dominicis, S., Rotko, P., Aaltonen, J., Waylen, K., & Langan, S. J. (2012). Enhancing flood resilience through improved risk communications. *Natural Hazards and Earth System Science*, *12*(7), 2271–2282. https://doi.org/10.5194/nhess-12-2271-2012

Pampapura Madali, N., Alsaid, M., & Hawamdeh, S. (2022). *Social media and people's perception of global warming during critical environmental events: the impact of misinformation through the lens of social noise* .

Sampson, N. R., Price, C. E., Kassem, J., Doan, J., & Hussein, J. (2019). "We're Just Sitting Ducks": Recurrent Household Flooding as An Underreported Environmental Health Threat in Detroit's Changing Climate. *International Journal of Environmental Research and Public Health* , *16*(1). https://doi.org/10.3390/ijerph16010006

Scott, K. K., & Errett, N. A. (2018). Content, Accessibility, and Dissemination of Disaster Information via Social Media during the 2016 Louisiana Floods. *Journal of Public Health Management and Practice*, *24*(4), 370–379. https://doi.org/10.1097/PHH.0000000000000708

Shoyama, K., Cui, Q., Hanashima, M., Sano, H., & Usuda, Y. (2021). Emergency flood detection using multiple information sources: Integrated analysis of natural hazard monitoring and social media data. *Science of the Total Environment*, *767*. https://doi.org/10.1016/j.scitotenv.2020.144371

State of Michigan. (2024). *State of Michigan Executive Branch*. https://www.michigan.gov/som/government/branches-of-government/executive-branch

Tabari, H. (2020). Climate change impact on flood and extreme precipitation increases with water availability. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-70816-2

Van Ootegem, L., & Verhofstadt, E. (2016). Well-being, life satisfaction and capabilities of flood disaster victims. *Environmental Impact Assessment Review*, *57*, 134–138. https://doi.org/10.1016/j.eiar.2015.12.001

Vo, B.-K. H., & Collier, N. (2013). Twitter Emotion Analysis in Earthquake Situations. *International Journal of Computational Linguistics and Applications*, *4*(1), 159–173.

Waite, T., Murray, V., & Baker, D. (2014). Carbon Monoxide Poisoning and Flooding: Changes in Risk Before, During and After Flooding Require Appropriate Public Health Interventions. *PLoS Currents*. https://doi.org/10.1371/currents.dis.2b2eb9e15f9b98278493880358448 7f1

Wang, Q., & Taylor, J. E.. (2015). *Process Map for Urban-Human Mobility and Civil Infrastructure Data Collection Using Geosocial Networking Platforms*. https://doi.org/10.1061/(ASCE)CP.1943-5487

Wolshon, P. B., Urbina, E. A., & Levitan, M. L. (2001). *National review of hurricane evacuation plans and policies*. Baton Rouge, LA: Louisiana State University Hurricane Center.

Zhou, S., Kan, P., Huang, Q., & Silbernagel, J. (2023). A guided latent Dirichlet allocation approach to investigate real-time latent topics of Twitter data during Hurricane Laura. *Journal of Information Science*, *49*(2), 465–479. https://doi.org/10.1177/01655515211007724

APPENDIX A. UMAP, HDBSCAN, and BERTopic Model Parameters

```
umap_model = umap.UMAP(n_neighbors= 10,
                       n_components= 10,
                       min_dist=0.0,
                       metric='cosine',
                       random_state= 632
                       )


hdbscan_model = HDBSCAN(min_cluster_size= 622,
                        min_samples=5,
                        metric='euclidean',
                        prediction_data=True)


topic_model = BERTopic(language="english",
                       min_topic_size = 622,
                       top_n_words = 10,
                       calculate_probabilities=False,
                       verbose=True,
                       umap_model=umap_model,
                       hdbscan_model=hdbscan_model)
```

Broadcasting Resilience: Developing Wildfire Readiness Through Student-Driven Radio Programs

José Azevedo
University of Porto
azevedo@letras.up.pt

Susana Neves
University of Porto
susanaalbernazneves@gmail.com

Ivone Neiva Santos
University of Porto
ivoneneivasantos@gmail.com

ABSTRACT

This paper introduces an innovative educational initiative focused on producing a community radio program to enhance students' understanding of wildfires. It creates an open dialogue platform for students to share ideas and concerns about wildfire preparedness, promoting an inclusive and participatory learning environment. A key component of this educational approach is establishing a 'readiness' framework to equip students with the necessary competencies to converse thoughtfully and effectively about the multifaceted narratives, dilemmas, and tensions inherent in dealing with significant wildfires. The project also emphasizes the importance of risk communication, aiming to improve students' abilities to convey information about wildfire risks clearly and accurately, catering to diverse audience needs. This approach is vital for enabling students to make informed decisions and take appropriate actions before, during, and after wildfire events, thus integrating scientific understanding and practical communication skills to foster proactive community engagement.

Keywords

Radio, youth engagement, preparedness, wildfires.

INTRODUCTION

Wildfire management is a complex and multifaceted challenge that requires a comprehensive approach to address effectively. The importance of changing attitudes, improving capacity, investing in community participation, and modifying stakeholder engagement procedures and policies is underscored by numerous studies, indicating that traditional methods of wildfire management are no longer sufficient in the face of increasing wildfire frequency and intensity due to climate change and other factors. Tedim et al. (2016) conceptualizes fire as a biophysical phenomenon and a fundamentally social process intricately tied to the human-natural systems in which it occurs. This perspective is echoed and expanded upon by Thompson et al. (2015) and Schultz and Moseley (2019), who call for systemic changes in wildfire management. They argue that addressing the current challenges of wildfires transcends the mere application of a set of theoretically sound procedures. Instead, it demands a localized understanding of wildfires, tailoring strategies to make territories less prone to wildfires and communities less vulnerable yet more resilient. This approach aligns with the goals of economic valorization, sustainable development, and the safeguarding of territorial resources.

Thompson specifically advocates for integrating risk management principles into decision-making processes, whereas Schultz and Moseley (2019) focus on the importance of collaborative governance to effectively tackle the complex challenges of wildfire management. Their research collectively underlines the necessity for a comprehensive strategy beyond traditional firefighting tactics. Such a strategy emphasizes a proactive

management stance, focusing on anticipating, assessing, and mitigating wildfire risks, underscoring the importance of a holistic plan that incorporates scientific understanding and community involvement to enhance resilience and safety.

From the assumption that improving societal readiness for wildfires requires a holistic approach to risk communication and attitude modification, it is clear that comprehensive methods are needed to successfully address the complex issues related to wildfire threats. This viewpoint emphasizes how crucial it is to incorporate many approaches and involve various stakeholders to foster a more resilient and knowledgeable community. In order to be effective, risk communication must go beyond the conventional one-way information transmission approaches and adopt interactive, participatory techniques that foster understanding and conversation. It is essential to leverage a range of technology and communication channels to reach different audiences. This will ensure that information about the risks posed by wildfires and how to prepare for them is clear, practical, and easily accessible.

The research presented here advocates a multidimensional strategy focused on three key pillars: a) the value of youth engagement and participation, b) establishing school-community learning collaborations, and c) emphasizing intergenerational capacity-building methods. Collectively, these pillars aim to enhance community resilience, encourage collaborative governance, and ensure a holistic and inclusive approach to wildfire risk communication and management.

YOUTH ENGAGEMENT AND PARTICIPATION

Promoting youth engagement and active involvement in society has become increasingly important in the past few decades (DFID-CSO, 2010). Participatory research programs have demonstrated their efficacy as interventions in promoting individual and community development. These initiatives operate under the premise that the current generation has valuable knowledge and should express their opinions firsthand, drawing from their own experiences (Kim, 2019, among others). According to the systematic review conducted by Anyon et al. (2018), youth participatory research programs have yielded significant outcomes in various domains, including social, interpersonal, cognitive, academic, and professional achievements, as well as leadership and personal agency.

Therefore, engaging youth in wildfire preparedness and environmental stewardship initiatives plays a vital role in building a resilient future, providing them with the necessary knowledge to grasp the complexities of wildfire risks and environmental conservation. Giving youth the tools to participate in prevention efforts and respond adeptly and effectively in emergencies instils a sense of responsibility and agency.

Moreover, engaging youth from an early age promotes a proactive mindset towards disaster risk management, emphasizing the importance of anticipation, planning, and prevention over a reactive stance (Pfefferbaum et al., 2018). This approach is crucial in preparing communities to respond to wildfires efficiently and engage actively in risk mitigation through thoughtful land management, community planning, and policy advocacy. The active participation of youth in these initiatives nurtures a culture of resilience, wherein the values, knowledge, and skills acquired are passed down and expanded upon, influencing peers, families, and future generations. Such a generational transfer solidifies a collective approach towards resilience, transforming how communities f ace wildfires and environmental challenges and ensuring a sustainable and secure future for all.

SCHOOL-COMMUNITY LEARNING PARTNERSHIPS

Through the utilization of educational institutions as primary hubs for community participation and resource sharing, these collaborations enable a comprehensive and inclusive educational experience that goes beyond the physical boundaries of the classroom, reaching families, local organizations, and various community members (Monroe et al., 2015; DePetris & Eames, 2017). This collaborative model enhances the learning ecosystem by spreading knowledge and skills about wildfire preparedness throughout the community (Paveglio & Edgeley, 2020). This fosters a shared feeling of responsibility and encourages collective action to reduce the dangers of wildfires.

These partnerships offer a distinctive platform for engaging and experiential learning activities that integrate theoretical knowledge with practical applications.

INTERGENERATIONAL APPROACHES IN CAPACITY BUILDING

The significance of intergenerational approaches in capacity building for wildfire preparedness cannot be overstated (Prior & Eriksen, 2013). These initiatives promote community resilience against wildfires by promoting communication and collaboration across multiple age groups, thereby using each generation's unique

capabilities and views. Elders in the community often possess a wealth of knowledge about local landscapes, traditional fire management practices, and historical patterns of wildfire activity. By sharing this expertise, they provide younger generations valuable insights that can supplement and amplify current scientific understanding and technical progress in wildfire management.

Moreover, involving youth in intergenerational learning initiatives empowers them to become proactive agents of change (Trujillo-Torres et al., 2023). Young people bring fresh perspectives, energy, and a natural aptitude for technology and innovation, which can be instrumental in developing new solutions for wildfire detection, monitoring, and communication.

INTERVENTION

We developed a school-based intervention, Forest FM, based on the foundational pillars described above. The following section will present the aims, procedures, and results.

Aims

In 2017, Portugal experienced devastating wildfires that ravaged over 539,000 hectares of land and claimed the lives of more than 100 individuals. These catastrophic events were partly attributed to the effects of climate change, leaving profound economic and environmental scars on the affected regions. It also highlighted the population's vulnerability in rural areas and underscored the need for measures to ensure assistance and increase community preparedness for the risks associated with climate change.

ForestFM is an educational initiative that emerged in response to the crisis of devastating wildfires, with a primary aim to educate and engage young people about the dangers of wildfires and the critical importance of preparedness. Central to the project's approach is establishing a community radio program by students, serving as a powerful platform to disseminate invaluable insights drawn from the experiences of local elders. Through this initiative, teenagers acquire knowledge about wildfire risks and actively preserve and honour traditional wisdom while gaining essential skills to access and effectively communicate vital information.

The implementation of ForestFM reflects a comprehensive approach to wildfire preparedness, integrating the principles of youth engagement, school-community collaboration, and intergenerational knowledge sharing. The project's immersive workshop on radio production, spanning multiple sessions, serves as a testament to this approach. In the initial training block, students are introduced to radio as a medium, alongside discussions on wildfire content and formats. This phase establishes a foundational understa nding and lays the groundwork for appreciating the significance of youth involvement in wildfire preparedness.

Subsequently, the workshop progresses to a deeper exploration of radio production techniques, emphasizing the use of sound and language to convey messages effectively. Significantly, students are sensitized to the importance of listening to and incorporating local knowledge and community perspectives into their content. This aspect underscores the value of school-community partnerships in fostering a more inclusive and informed approach to wildfire communication and preparedness.

In the final phase of the workshop, students actively apply their newfound knowledge and skills to produce radio content explicitly focused on wildfire preparedness. This culmination reflects the collaborative efforts of youth, educators, and local community members, embodying the intergenerational exchange of knowledge and the collective commitment to wildfire resilience.

Procedures

The study included students from three high schools in central Portugal's high-risk wildfire zones. The methodology focused on an intervention in which students created radio content. The study aims to assess the impact of the intervention on students' knowledge, awareness, and information -seeking behaviours related to wildfires by conducting surveys before and after the workshop. This comprehensive methodology allowed for a complete analysis of the intervention's effects and offered helpful information for designing future wildfire education programs aimed at teenagers in high-risk regions.

The first survey aimed to understand local adolescents' collective knowledge and awareness levels by targeting a more significant cohort of students across the three schools (N=234). This method yielded insights that extended beyond the workshop attendees. The subsequent survey was only given to the workshop participants (N= 49), allowing for a more targeted examination of how the intervention affected their awareness and attitudes concerning wildfires.

Both surveys had a specific area to evaluate where participants obtained information about wildfires. Participants evaluated a set of statements of information-seeking behaviour using a 5-point scale ranging from 1 (total disagreement) to 5 (complete agreement). However, for the purpose of this paper, we specifically focus on the effects of the intervention on social conversation related to wildfires. This includes examining the frequency of discussions with family and friends as well as self-perception of wildfires knowledge

Results

In this section, we present the findings from a pre-post survey designed to assess changes in participants' engagement in social conversations about wildfires with family and friends and their self-perception of knowledge on the subject. Our analysis focuses on three key questions: the frequency of discussions regarding wildfire preparedness within their networks, the depth and quality of these conversations, and the participants' self-assessed understanding of wildfires before and after engaging with the radio program. These aspects are crucial for evaluating the program's impact on fostering a culture of preparedness and awareness through informed dialogue and self-reflection. By examining these indicators, we aim to shed light on the effectiveness of the radio program in enhancing not just awareness but meaningful communication and self-perceived knowledge about wildfires among community members.

Table 1. Levels of Agreement on Frequency of Wildfire Discussions with Family and Friends on Self-perception of Preparedness (%)

	Strongly Disagree		**Somewhat Disagree**		**Neutral**		**Somewhat Agree**		**Strongly Agree**	
	Pre	**Post**	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**
I frequently talk about wildfires with my family members	21.8	4,1	24.8	16.3	32.1	40.8	18.4	32,7	3.0	6.1
I usually talk about wildfires with my friends	28.6	2.0	23.9	34,7	25.8	26.1	17.1	24.5	4.3	12.2
I have sufficient knowledge of what to do in the face of wildfires	2,6	0.0	15.5	4.1	25.3	40.8	17.2	49.0	9.4	6.1

Table1 presents findings from a pre-post survey designed to assess changes in behaviours and perceptions related to wildfire discussions and self-assessed knowledge following an educational intervention. Here are some key points developed further from the table:

1. **Increased Conversations about Wildfires**: There is a notable shift in the frequency of discussions participants have about wildfires with their family and friends. Before the intervention, a smaller percentage of respondents reported frequently discussing wildfires with their family (3.0% strongly agree) and friends (4.3% strongly agree). Post-intervention, these numbers increased to 6.1% for family discussions and 12.2% for discussions with friends, indicating the intervention's effectiveness in encouraging more conversations about wildfire preparedness.
2. **Change in Discussion Dynamics**: The pre-survey data showed that a significant portion of participants were either neutral or disagreed with the statement about frequently discussing wildfires with family and friends. Post-intervention, there is a shift towards agreement, especially in the "somewhat agree" category for family discussions, which almost doubled from 18.4% to 32.7%. This suggests a positive change in participants' willingness or opportunity to engage in conversations about wildfires.
3. **Self-Assessed Knowledge Improvement**: The most striking change is observed in the participants' self-assessment of their knowledge regarding what to do in the face of wildfires. Initially, only 9.4%

strongly agreed with having sufficient knowledge, but post-intervention, there was a significant increase in the "somewhat agree" category from 17.2% to 49.0%. This reflects a substantial improvement in confidence among participants about their understanding of wildfire preparedness. However, the "strongly agree" category decreased slightly, suggesting room for further enhancement in perceived knowledge depth.

4. **Decrease in Neutral Responses**: Post-intervention, there is a noticeable decrease in neutral responses across all questions, most significantly in the self-assessed knowledge about wildfires, dropping from 25.3% to 40.8%. This reduction suggests that the intervention helped clarify participants' uncertainties or ambiguities regarding wildfires, pushing them towards a more decisive stance on their knowledge and discussion frequency.
5. **Reduction in Disagreement**: There is a general decrease in the percentages of "strongly disagree" and "somewhat disagree" responses from pre- to post, particularly in conversations about wildfires with family. This trend suggests that reservations or reluctance to engage in these discussions have been mitigated, possibly due to increased knowledge and awareness.

CONCLUSION

A comprehensive strategy for wildfire management necessitates understanding the fire as a biophysical and social process, as Tedim et al. (2016) suggest, and a proactive and inclusive approach to learning and adaptation. This strategy should emphasize youth engagement, intergenerational capacity building, and school-community partnerships as central pillars. By integrating these concepts, communities can foster a culture of continuous learning and adaptation to the dynamic nature of wildfire threats.

Regular evaluation of community participation efforts, risk communication campaigns, and preparedness operations becomes crucial in identifying areas for improvement. Learning from past wildfires and adjusting strategies based on new information and changing conditions are essential to enhance societal preparedness. Moreover, involving young people and creating opportunities for intergenerational learning can lead to more innovative solutions and a deeper understanding of wildfire risks.

School-community partnerships can play a pivotal role in this context, bridging knowledge and action. Through educational programs focusing on wildfire preparedness, risk communication, and resilience building, these partnerships can empower communities to be better prepared for, respond to, and recover from wildfires. Such a comprehensive strategy, grounded in community involvement and intergenerational collaboration, enhances resilience to wildfires and helps safeguard communities and conserve ecosystems against the backdrop of evolving climatic conditions and increasing wildfire threats.

The educational initiative Forest FM aspires to create a more informed and resilient generation of students through its focus on open dialogue, readiness framework, and the strategic incorporation of risk communication. By developing their understanding of and ability to communicate about wildfires, the project supports the broader goal of enhancing community preparedness and resilience in the face of natural disasters.

The data from the Forest FM project, which involved the creation of radio content focused on wildfires, reveal a significant increase in discussions about wildfires among participants with their families and friends, as well as an improvement in their self-assessed understanding of how to respond to these events. The positive shift in responses post-intervention highlights the effectiveness of targeted educational initiatives in cultivating a more informed and proactive community regarding wildfire awareness and preparedness. This project equipped individuals with the knowledge and skills to address wildfire concerns and contributed to a broader cultural shift towards environmental sustainability and resilience.

The effort highlights the need for education to develop a well-informed and adaptable society to face wildfire concerns by promoting community-wide involvement in preparedness measures. Schools play a crucial role in developing resilient, knowledgeable, and cohesive communities by integrating these programs into their curricula. This all-encompassing educational method promotes a broad cultural shift towards environmental sustainability and resilience, preparing communities to be more resilient and proactiv e in dealing with environmental challenges in the future.

ACKNOWLEDGMENTS

ForestFM (PCIF/AGT/0087/2019) is a project funded by the Portuguese Foundation for Science and Technology (FCT) and seeks to promote the engagement of young people in promoting attitudes and behaviors favorable to wildfire prevention.

REFERENCES

Anyon, Y., Bender, K., Kennedy, H., & Dechants, J. (2018). A Systematic Review of Youth Participatory Action Research (YPAR) in the United States: Methodologies, Youth Outcomes, and Future Directions. *Health Education & Behavior: The Official Publication of the Society for Public Health Education*, 45(6), 865–878. https://doi.org/10.1177/1090198118769357

DePetris, T., & Eames, C. (2017). A collaborative community education model: Developing effective school-community partnerships. *Australian Journal of Environmental Education, 33* (3), 171–188. https://doi.org/10.1017/aee.2017.26

DFID-CSO Youth Working Group. (2010). *Youth participation in development: A guide for development agencies and policy makers*. https://www.youthpolicy.org/library/documents/youth-participation-in-development-a-guide-for-development-agencies-and-policy-makers/

Kim, J. (2019). Youth involvement in Participatory Action Research (PAR): Challenges and barriers. *Critical Social Work, 17*(1), 38–53. https://doi.org/10.22329/csw.v17i1.5891

Monroe, M. C., Ballard, H. L., Oxarart, A., Sturtevant, V. E., Jakes, P. J., & Evans, E. R. (201 5). Agencies, educators, communities and wildfire: Partnerships to enhance environmental education for youth. *Environmental Education Research, 22*(8), 1098-1114. http://dx.doi.org/10.1080/13504622.2015.1057555

Paveglio, T. B., Edgeley, C. M. (2020). Fire Adapted Community. In: Manzello, S. (eds) *Encyclopedia of Wildfires and Wildland-Urban Interface (WUI) Fires*. Springer, Cham. https://doi.org/10.1007/978-3-319-51727-8_114-1

Pfefferbaum, B., Pfefferbaum, R. L., & Van Horn, R. L. (2018). Involving children in disaster risk reduction: the importance of participation. *European Journal of Psychotraumatology*, *9*(sup2). https://doi.org/10.1080/20008198.2018.1425577

Prior, T., & Eriksen, C. (2013). Wildfire preparedness, community cohesion and social–ecological systems. *Global environmental change, 23*(6), 1575-1586. https://doi.org/10.1016/j.gloenvcha.2013.09.016

Schultz, C. A., & Moseley, C. (2019). Collaborations and capacities to transform fire management. *Science*, *366*(6461), 38–40. https://www.science.org/doi/10.1126/science.aay3727

Tedim, F., Leone, V., Xanthopoulos, G. (2016). A wildfire risk management concept based on a social-ecological approach in the European Union: Fire Smart Territory. *International Journal of Disaster Risk Reduction, 18*, 138–153. https://doi.org/10.1016/j.ijdrr.2016.06.005

Thompson, M., Dunn, C., Calkin, D. (2015). Wildfires: Systemic changes required. *Science*. *350*(6263), 920–920. https://www.science.org/doi/10.1126/science.350.6263.920-b

Trujillo-Torres, J. M., Aznar-Díaz, I., Cáceres-Reche, M. P., Mentado-Labao, T., & Barrera-Corominas, A. (2023). Intergenerational Learning and Its Impact on the Improvement of Educational Processes. *Education Sciences, 13*(10), 1019. https://doi.org/10.3390/educsci13101019



Social Media for Crisis Management

Track Chairs

Shelby Knudsen, Trilateral Research

Su Anson, Trilateral Research

Jeannette Annies, University of Stuttgart

Semantically Enriched Cross-Lingual Sentence Embeddings for Crisis-related Social Media Texts

Rabindra Lamsal*
The University of Melbourne
rlamsal@student.unimelb.edu.au

Maria Rodriguez Read
The University of Melbourne
maria.read@unimelb.edu.au

Shanika Karunasekera
The University of Melbourne
karus@unimelb.edu.au

ABSTRACT

Tasks such as semantic search and clustering on crisis-related social media texts enhance our comprehension of crisis discourse, aiding decision-making and targeted interventions. Pre-trained language models have advanced performance in crisis informatics, but their contextual embeddings lack semantic meaningfulness. Although the CrisisTransformers family includes a sentence encoder to address the semanticity issue, it remains monolingual, processing only English texts. Furthermore, employing separate models for different languages leads to embeddings in distinct vector spaces, introducing challenges when comparing semantic similarities between multi-lingual texts. Therefore, we propose multi-lingual sentence encoders (CT-XLMR-SE and CT-mBERT-SE) that embed crisis-related social media texts for over 50 languages, such that texts with similar meanings are in close proximity within the same vector space, irrespective of language diversity. Results in sentence encoding and sentence matching tasks are promising, suggesting these models could serve as robust baselines when embedding multi-lingual crisis-related social media texts. The models are publicly available at: https://huggingface.co/crisistransformers.

Keywords

Sentence Encoders, Embedding Models, Cross-lingual Vector Space, Multi-lingual Embeddings, CrisisTransformers

INTRODUCTION

In times of crisis, social media platforms such as Facebook and Twitter are critical channels for information sharing and communication (Imran, Castillo, Diaz, et al. 2015; Simon et al. 2015; Lamsal, Harwood, et al. 2022a). These platforms help promptly disseminate essential information, whether it is related to wildfires, earthquakes, hurricanes, floods, epidemics, etc (Starbird and Palen 2010; Thomson et al. 2012; Alam, Ofli, Imran, and Aupetit 2018; Pourebrahim et al. 2019; Lamsal, Read, et al. 2023a). They serve as central information hubs for both the general population and emergency responders, providing updates on unfolding situations (Sarcevic et al. 2012). Social media also assists in allowing individuals to seek and offer assistance while coordinating relief efforts (Purohit et al. 2014). The extensive amount of user-generated content on these platforms is a valuable historical and real-time data source. However, the challenges of analyzing and understanding crisis-related social media texts arise from their sheer volume (Stieglitz et al. 2018) and linguistic complexities. As the number of conversations exponentially increases during a crisis, automated analysis becomes necessary to comprehend the situation at the ground level. Such conversations contain situational information (Hughes and Palen 2009; Vieweg et al. 2010; Vieweg 2012) about the affected population, damages, injuries, casualties, rescue and volunteering efforts, etc. (Imran, Mitra, et al. 2016), shared by populations, whether directly or indirectly impacted (Imran, Castillo, Diaz, et al. 2015; Lamsal, Harwood, et al. 2022a).

*corresponding author

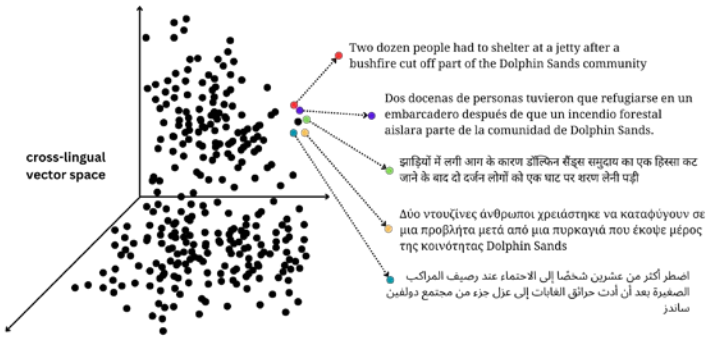


Figure 1. An illustration of a cross-lingual vector space for crisis-related social media texts.

The majority of analyses in processing crisis-related social media texts involve tasks like text classification (Ashktorab et al. 2014; Caragea et al. 2011; Imran, Elbassuoni, et al. 2013a; Li et al. 2018), semantic search (Dutt et al. 2019), and clustering (Ashktorab et al. 2014; Curiskis et al. 2020; Lamsal, Harwood, et al. 2022a). The advancement in these areas plays a key role in enhancing our comprehension of crisis discourse to facilitate informed decision-making processes and assist in the development of targeted interventions and communication strategies. This holds true not only during the course of an unfolding crisis but also retrospectively, drawing insights from historical events. The knowledge extracted from previous crises significantly contributes to the formulation of effective and efficient response strategies for similar future crisis scenarios, thus strengthen preparedness and response capabilities for future challenges.

Transformer-based (Vaswani et al. 2017) pre-trained language models like BERT (Devlin et al. 2018) and RoBERTa (Liu et al. 2019) have significantly advanced performance in numerous NLP tasks across domains including crisis informatics. Moreover, the recently introduced CrisisTransformers (Lamsal, Read, et al. 2023b), an ensemble of pre-trained models trained on a corpus of over 15 billion word tokens from more than 30 crisis events, has further improved the state-of-the-art. However, the contextual embeddings provided by these pre-trained models lack semantic meaningfulness[1] and perform worse than averaging GloVe embeddings (Reimers and Gurevych 2019). For semantic search and clustering tasks, it is crucial to have sentence embeddings with semantic richness. While the CrisisTransformers family does offer a sentence encoder for the task, improving upon the Sentence Transformers (Reimers and Gurevych 2019), it remains monolingual, processing only English-language social media texts. Any geographical region can have a diverse linguistic population, whether it's a county, city, state, or country. Social media platforms, therefore, can be flooded with posts in various languages during the same crisis event. Analyzing solely the texts in a specific language increases the risk of overlooking essential information available in texts shared in other languages. Also, employing separate models for different languages leads to embeddings in distinct vector spaces. This discrepancy poses a challenge when comparing the semantic similarities between sentences in various languages.

Adopting a single embedding space is critical to addressing these challenges and enhancing the effectiveness of crisis informatics. Such an encoder would facilitate mapping texts from different languages with similar meanings to close proximity within the same vector space, as illustrated in Figure 1. This approach ensures that semantically related content in diverse languages can be effectively processed, enabling a more comprehensive and nuanced understanding of crisis-related information irrespective of linguistic diversity. Therefore, we address this need by introducing the first-ever multi-lingual sentence encoders for embedding crisis-related social media texts. In general, this study contributes the following to the existing crisis informatics literature:

- We introduce two multi-lingual sentence encoders (CT-XLMR-SE[2] and CT-mBERT-SE[3]) that embed crisis-related social media texts with semantic richness for 52 languages: Albanian, Arabic, Armenian, Bulgarian, Catalan, Chinese (Simplified), Chinese (Traditional), Croatian, Czech, Danish, Dutch, Estonian, Finnish, French, French (Canada), Galician, Georgian, German, Greek, Gujarati, Hebrew, Hindi, Hungarian, Indonesian, Italian, Japanese, Korean, Kurdish (Sorani), Latvian, Lithuanian, Macedonian, Malay, Marathi, Mongolian, Myanmar (Burmese), Norwegian, Persian, Polish, Portuguese, Portuguese (Brazil), Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swedish, Thai, Turkish, Ukrainian, Urdu, and Vietnamese.

[1] Semantically similar sentences are embedded together in a vector space, and those embeddings can be compared using similarity measures such as cosine similarity.

[2] https://huggingface.co/crisistransformers/CT-XLMR-SE

[3] https://huggingface.co/crisistransformers/CT-mBERT-SE

- We publicly release the sentence encoders, making them easily accessible for integration with the Transformers library (Wolf et al. 2020). We anticipate that these sentence encoders will serve as robust baselines for tasks that involve embedding multi-lingual crisis-related social media texts.

RELATED WORK

Semantically rich embeddings position similar sentences closely together in a vector space. Learning such embeddings is an extensively explored area, briefly discussed in this section. (Kiros et al. 2015) trained an encoder-decoder model to reconstruct the neighbouring sentences of an encoded sequence. This training aimed to map sentences with similar semantic properties to comparable vector representations. (Conneau, Kiela, et al. 2017) introduced a siamese Bidirectional Long Short-Term Memory (BiLSTM) network, incorporating max-pooling on the Stanford Natural Language Inference (SNLI) dataset. This approach surpassed previous unsupervised methods (Kiros et al. 2015; Hill et al. 2016). (Cer et al. 2018) extended unsupervised learning by training a transformer network on the SNLI dataset. Furthermore, (Yang, Yuan, et al. 2018) proposed an unsupervised learning method for sentence-level semantic similarity based on conversational data.

With the introduction of BERT in 2018, the unsupervised training component was replaced with pre-trained models in designing sentence encoders. (Reimers and Gurevych 2019) fine-tuned BERT using siamese and triplet networks on the SNLI and Multi-Genre Natural Language Inference (MultiNLI) datasets. The fine-tuning involved a softmax classifier with "contradiction", "entailment", and "neutral" labels. Similarly, (Gao et al. 2021) proposed SimCSE, a contrastive approach to fine-tune pre-trained models using natural language inference datasets, utilizing "contradiction" pairs as hard negatives. Building on this, (Reimers and Gurevych 2019) fine-tuned multiple pre-trained models using over 1 billion sentence pairs and released the second version of their models. Following the approach of utilizing "contradiction" pairs as hard negatives, as proposed in (Gao et al. 2021), (Lamsal, Read, et al. 2023b) replaced pre-trained models such as BERT and RoBERTa, which were trained on texts from broad and general domains with crisis-specific pre-trained models. Their domain-specific models, publicly released as CrisisTransformers, were trained on an extensive corpus containing over 15 billion word tokens from 30+ crisis events, including disease outbreaks, natural disasters, protests and activism, conflicts, civil war, etc. Performance improvement of >17% was reported compared to Sentence Transformers (Reimers and Gurevych 2019) in sentence encoding tasks for crisis-related social media texts. However, their sentence encoder is monolingual, processing only English language texts.

Multiple approaches have been proposed in the literature to train multi-lingual sentence embeddings. (Artetxe and Schwenk 2019) trained an encoder-decoder network on parallel corpora with a translation task and used output from the encoder as the sentence embedding. While this approach is effective in identifying exact translations across various languages, its performance diminishes when evaluating the similarity of sentences that are not exact translations. (Yang, Cer, et al. 2019) learnt multi-lingual sentence embeddings using a multi-task setup on the SNLI dataset and millions of question-answer pairs. They used a translation ranking task to align cross-lingual vector spaces. However, this approach is susceptible to catastrophic interference and has significant computational overhead. In addressing these issues, (Reimers and Gurevych 2020) proposed a training approach (Sentence Transformers) to map a translated sentence to the same location in the vector space as the original sentence. In this way, a multi-lingual model (student model) can be trained on translation pairs to mimic a mono-lingual model (teacher model) and learn a common vector space for multiple languages. In this study, we consider Sentence Transformer's top-performing model *all-mpnet-base-v2* as a baseline.

MATERIALS AND METHODS

Training Architecture

We utilized CrisisTransformers' sentence encoder (Lamsal, Read, et al. 2023b) as a teacher model and extended it to develop multi-lingual models with a student-teacher training architecture (Reimers and Gurevych 2020). The training network is illustrated in Figure 2. We considered XLM-R (Conneau, Khandelwal, et al. 2019) and mBERT (Devlin et al. 2018) as student models. XLM-R is the multi-lingual version of RoBERTa, trained on 2.5TB of CommonCrawl data containing 100 languages. Similarly, mBERT is the multi-lingual version of BERT, trained on Wikipedia data for 104 languages.

Training: For a dataset $D = ((en_1,t_1), (en_2,t_2), ..., (en_n,t_n))$, where en_i are English sentences and t_i are respective translated sentences (can be any language), we trained a network with student model S and teacher model CT while minimizing the following training objective for a given mini-batch B:

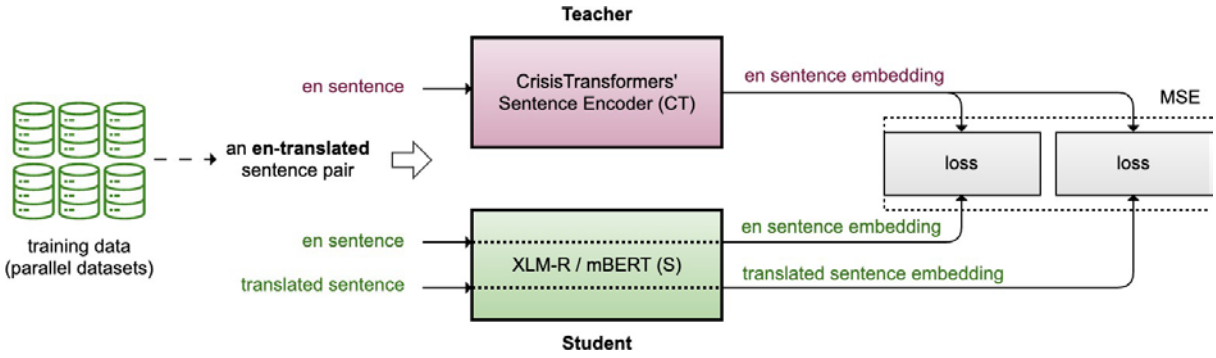


Figure 2. The student-teacher training architecture.

$$\frac{1}{|B|}\sum_{j\in B}\left[(\mathrm{CT}(en_j)-\mathrm{S}(en_j))^2+(\mathrm{CT}(en_j)-\mathrm{S}(t_j))^2\right] \quad (1)$$

With the above training objective, trained on dataset D we aim to obtain S such that $\mathrm{S}(en_i) \approx \mathrm{CT}(en_i)$ and $\mathrm{S}(t_i) \approx \mathrm{CT}(en_i)$. S will be our multi-lingual model. We train XLM-R and mBERT in separate networks with common teacher CT , and their final checkpoints are named CT-XLMR-SE and CT-mBERT-SE (where *SE* stands for *sentence encoder*). The training was done on an NVIDIA A100 GPU (80GB) for a maximum of 20 epochs, and 20k steps were used to warm up the learning rate with AdamW as an optimizer. Mixed precision training was done to improve the training time. We implemented the following training setting: maximum sequence length: 128, batch size: 64, maximum sentence per training file: 500k, and learning rate: 2e-5. The training of both networks in parallel finished in approximately 30 days.

Training Data

The training data D comprised multiple publicly available parallel datasets merged together. For each language, the maximum number of sentence pairs considered for training from any single dataset was 500k. Some samples in certain datasets had a single entry, either a sentence in English or any other language; in such cases, those samples were ignored. In total, the training data included over 128 million sentence pairs. The released versions of all the datasets already had the texts pre-processed. Some example pairs from the training data are provided in Table 1. The following parallel datasets were considered in this study:

- **Europarl** (Koehn 2005): Sentences extracted from the proceedings of the European Parliament.
- **GlobalVoices** (Tiedemann 2012): News stories collected from the website globalvoices.org.
- **JW300** (Agic and Vulic 2019): Publications crawled from the website jw.org.
- **MUSE** (Conneau, Lample, et al. 2017): Ground-truth bilingual dictionaries.
- **News-Commentary** (Barrault et al. 2019): News commentaries collected from the website Project Syndicate, provided by WMT for shared tasks.
- **OpenSubtitles** (Lison and Tiedemann 2016): Large-scale collection of movie and TV subtitles.
- **Tatoeba**[4]: Large-scale collection of sentences and their translations.
- **TED2020** (Reimers and Gurevych 2020): Translated subtitles of over 4,000 TED talks.
- **WikiMatrix** (Schwenk et al. 2019): Parallel sentences from Wikipedia's contents.
- **WikiTitles** (Reimers and Gurevych 2020): Cross-lingual Wikipedia titles extracted from Wikipedia dumps.

4https://tatoeba.org/

Table 1. Example sentence pairs.

english sentence (en)	translated sentence (t)
1917 also saw the last issuance of 2 kroner coins.	[Portuguese] 1917 tambe´m viu a última emissa˜o de moedas de 2 kroner.
Resumption of the session.	[Spanish] Reanudacio´n del per´ıodo de sesiones.
And then we can ask questions, real questions, questions like, what's the best life insurance policy to get? – real questions that people have in their everyday lives.	[German] Und dann können wir Fragen stellen, echte Fragen, Fragen wie: Welche ist die beste Lebensversicherung? – echte Fragen, die die Leute im ta¨glichen Leben haben.
Then they surrendered and so did the 4 other ships, one after another.	[Dutch] Toen gaven ze zich over, evenals de vier andere schepen, de een na de ander.
Declaration of Fundamental Rights and Freedoms adopted by the Parliament.	[French] De´claration des droits et liberte´s fondamentales adopte´ par le Parlement.
Most economists failed to foresee the economic dynamics that actually led to the crisis, because they failed to pay enough attention to the rapid increase in US total debt.	[Italian] La maggior parte degli economisti non riusc`ı a prevedere le dinamiche economiche che condussero davvero alla crisi, perche´ non presto` sufficiente attenzione al rapido incremento del debito complessivo americano.

Evaluations

Sentence Encoding Tasks

We implemented the sentence encoding task described in (Lamsal, Read, et al. 2023b) to evaluate the embeddings generated by students compared to those by the teacher. This assessment measures how closely the students have replicated the teacher's vector space. For a crisis-related labelled dataset, we calculated the weighted average cosine similarity among tweets within individual classes as below:

$$D_{\text{avg}} = \sum_{k=1}^{K} \hat{w_k} \cdot \frac{1}{|\{i : y_i = c_k\}|} \sum_{i:y=c_k} \text{similarity}(\mathbf{e}_i, \mathbf{e}_j) \quad (2)$$

where, K is the number of classes in a dataset, $\hat{w}_k$ is the normalized class weight, similarity$(\mathbf{e}_i, \mathbf{e}_j)$ computes the cosine similarity between embeddings $\mathbf{e}_i$ and $\mathbf{e}_j$, where $\mathbf{e}_j$ is a crisis-text from same class as $\mathbf{e}_i$.

Datasets: For this task, we utilized 18 publicly available human-labelled crisis-related texts datasets, all of which are in English. The following datasets were considered: CrisisMMD (Alam, Ofli, and Imran 2018), CrisisLex (Olteanu et al. 2014), AIDR (Imran, Castillo, Lucas, et al. 2014), ISCRAM2013 (Imran, Elbassuoni, et al. 2013a), SWDM2013 (Imran, Elbassuoni, et al. 2013b), CrisisNLP (Imran, Mitra, et al. 2016), COVID-19 stance (Poddar et al. 2022), SAD Stressor (Mauriello et al. 2021), SAD Stress (Mauriello et al. 2021), SAD COVID (Mauriello et al. 2021), LocBERT (Lamsal, Harwood, et al. 2022b), HMC (a) (Biddle et al. 2020), Vaccination opinions (Cotfas et al. 2021), HMC (b) (Biddle et al. 2020), PHM (Karisani and Agichtein 2018), COVID-19 patients and COVID-19 contacts (Klein et al. 2021), and ANTiVax (Hayawi et al. 2022).

Sentence Matching Tasks

Next, we assessed the multi-lingual aspect of the student models. We considered the test sets provided in TED2020 (Reimers and Gurevych 2020) to evaluate our student models. Each test set comprised 1,000 sentence pairs. For every test set and each student, we performed sentence matching tasks as follows:

Let E be the set of English sentences and T be the set of translated sentences. We have n sentence pairs, where $E = \{en_1, en_2, ..., en_n\}$ and $T = \{t_1, t_2, ..., t_n\}$. For each i (where $1 \leq i \leq n$), we compute the embeddings $S(en_i)$ and $S(t_i)$ for the English and translated sentences, respectively.

English to translated language evaluation: For each i, assess if t_i has the highest cosine similarity with en_i compared to all other translated sentences in the set T:

$$\text{en-t-match}_i = \begin{cases} 1 & \text{similarity}(S(en_i), S(t_i)) \geq \text{similarity}(S(en_i), S(t_j)) \text{ for all } t_j \neq t_i \\ 0 & \text{else} \end{cases} \quad (3)$$

Table 2. Results (weighted average cosine similarity) from sentence encoding tasks.

Dataset	SBERT	CT	CT-XLMR-SE	CT-mBERT-SE
CrisisMMD (Alam, Ofli, and Imran 2018)	0.6103	0.7191	0.6772	0.6724
CrisisLex (Olteanu et al. 2014)	0.6809	0.8070	0.7705	0.7734
AIDR (Imran, Castillo, Lucas, et al. 2014)	0.5407	0.6128	0.5586	0.5644
ISCRAM2013 (Imran, Elbassuoni, et al. 2013a)	0.5632	0.6872	0.6545	0.6546
SWDM2013 (Imran, Elbassuoni, et al. 2013b)	0.4870	0.6744	0.6312	0.6314
CrisisNLP (Imran, Mitra, et al. 2016)	0.5698	0.6960	0.6390	0.6446
COVID-19 stance (Poddar et al. 2022)	0.6019	0.8742	0.8444	0.8451
SAD Stressor (Mauriello et al. 2021)	0.7528	0.7027	0.6818	0.6864
SAD Stress (Mauriello et al. 2021)	0.8749	0.9021	0.8961	0.8974
SAD COVID (Mauriello et al. 2021)	0.6674	0.7167	0.7134	0.7147
LocBERT (Lamsal, Harwood, et al. 2022b)	0.7297	0.8982	0.8750	0.8834
HMC (a) (Biddle et al. 2020)	0.8211	0.8342	0.8101	0.8138
Vaccination opinions (Cotfas et al. 2021)	0.4939	0.7874	0.7543	0.7633
HMC (b) (Biddle et al. 2020)	0.7259	0.7762	0.7336	0.7379
PHM (Karisani and Agichtein 2018)	0.7579	0.7708	0.7478	0.7537
COVID-19 patients (Klein et al. 2021)	0.5320	0.6549	0.7054	0.7082
COVID-19 contacts (Klein et al. 2021)	0.5192	0.5874	0.6034	0.5984
ANTiVax (Hayawi et al. 2022)	0.5450	0.7722	0.7631	0.7614
Average	0.6374	0.7485	0.7255	0.7280

Table 3. Performance of vanilla XLM-R and mBERT on sentence encoding tasks.

	XLM-R	mBERT
Average	0.0032	0.3581

Translated language to English evaluation: For each i, assess if en_i has the highest cosine similarity with t_i compared to all other English sentences in the set E:

$$\text{t-en-match}_i = \begin{cases} 1 & \text{similarity}(S(t_i), S(en_i)) \geq \text{similarity}(S(t_i), S(en_j)) \text{ for all } en_j \neq en_i \\ 0 & \text{else} \end{cases} \quad (4)$$

In Equations 3 and 4, "1" denotes a correct match, and "0" an incorrect match. The accuracy for each language is calculated by dividing the total number of correct matches by the number of sentence pairs in the test set.

Sentence Embeddings

For generating sentence embeddings, we performed mean pooling on the token embeddings of an input sentence while considering the attention mask. The following pre-processing steps were performed to each input sequence: (i) replacing URLs with the HTTPURL token, (ii) substituting mentions with the @MENTION token, (iii) decoding HTML entities, (iv) eliminating newline characters and unnecessary whitespaces, (v) fixing text encoding issues for consistency, and (vi) replacing emojis with their corresponding textual representations. A tweet was considered one input sequence with the maximum sequence length set to 128.

RESULTS AND DISCUSSIONS

As discussed earlier, we implemented the student-teacher architecture and used a collection of parallel datasets, totalling over 128 million sentence pairs, as training data to train multi-lingual sentence encoders based on XLM-R and mBERT. The final checkpoints of the sentence encoders (CT-XLMR-SE and CT-mBERT-SE) were evaluated to assess: (i) how well they have learned the teacher's vector space and (ii) their multi-lingual capacity. The results from sentence encoding tasks are provided in Table 2 and Table 3, and those from sentence matching tasks are summarized in Table 4.

Table 4. Results (accuracy) from sentence matching tasks. For each language, the best average score is highlighted and equal scores across models are underlined. `en-t`: English to translated language and `t-en` is translated language to English.

	CT-XLMR-SE			CT-mBERT-SE		
Language	**en-t acc.**	**t-en acc.**	**avg.**	**en-t acc.**	**t-en acc.**	**avg.**
Albanian	0.9420	0.9480	0.9450	0.9510	0.9500	**0.9505**
Arabic	0.9430	0.9440	0.9435	0.9420	0.9380	**0.9400**
Armenian	0.9200	0.9100	**0.9150**	0.9170	0.9050	0.9110
Bulgarian	0.9670	0.9660	**0.9665**	0.9650	0.9640	0.9645
Catalan	0.9600	0.9610	**0.9605**	0.9520	0.9530	0.9525
Chinese (Simplified)	0.9420	0.9430	**0.9425**	0.9410	0.9280	0.9345
Chinese (Traditional)	0.9420	0.9280	0.9350	0.9460	0.9260	**0.9360**
Croatian	0.9710	0.9710	0.9710	0.9750	0.9700	**0.9725**
Czech	0.9610	0.9590	0.9600	0.9630	0.9630	**0.9630**
Danish	0.9640	0.9640	**0.9640**	0.9610	0.9650	0.9630
Dutch	0.9760	0.9720	**0.9740**	0.9760	0.9700	0.9730
Estonian	0.9340	0.9340	**0.9340**	0.9270	0.9220	0.9245
Finnish	0.9390	0.9390	0.9390	0.9430	0.9370	**0.9400**
French	0.9710	0.9680	**0.9695**	0.9730	0.9650	0.9690
French (Canada)	0.9530	0.9490	0.9510	0.9570	0.9460	**0.9515**
Galician	0.9630	0.9640	0.9635	0.9650	0.9620	0.9635
Georgian	0.8890	0.9000	0.8945	0.9130	0.9010	**0.9070**
German	0.9760	0.9710	**0.9735**	0.9740	0.9700	0.9720
Greek	0.9700	0.9700	**0.9700**	0.9660	0.9600	0.9630
Gujarati	0.8600	0.8650	0.8625	0.8990	0.8860	**0.8925**
Hebrew	0.9590	0.9610	**0.9600**	0.9580	0.9610	0.9595
Hindi	0.9060	0.9080	**0.9070**	0.8990	0.9010	0.9000
Hungarian	0.9620	0.9580	**0.9600**	0.9600	0.9550	0.9575
Indonesian	0.9740	0.9620	**0.9680**	0.9730	0.9610	0.9670
Italian	0.9640	0.9600	0.9620	0.9650	0.9610	**0.9630**
Japanese	0.9240	0.9150	**0.9195**	0.9240	0.9140	0.9190
Korean	0.9450	0.9420	**0.9435**	0.9350	0.9310	0.9330
Kurdish (Sorani)	0.8940	0.8840	**0.8890**	0.3460	0.2650	0.3055
Latvian	0.9560	0.9550	**0.9555**	0.9550	0.9520	0.9535
Lithuanian	0.9420	0.9310	**0.9365**	0.9310	0.9330	0.9320
Macedonian	0.9790	0.9820	**0.9805**	0.9790	0.9810	0.9800
Malay	0.9190	0.9190	**0.9190**	0.9120	0.9100	0.9110
Marathi	0.8300	0.8340	0.8320	0.8680	0.8580	**0.8630**
Mongolian	0.8430	0.8300	0.8365	0.8730	0.8670	**0.8700**
Myanmar (Burmese)	0.9200	0.9070	**0.9135**	0.9110	0.9010	0.9060
Norwegian	0.9470	0.9460	0.9465	0.9520	0.9470	**0.9495**
Persian	0.9290	0.9280	**0.9285**	0.9230	0.9250	0.9240
Polish	0.9560	0.9460	0.9510	0.9560	0.9490	**0.9525**
Portuguese	0.9700	0.9780	**0.9740**	0.9700	0.9770	0.9735
Portuguese (Brazil)	0.9780	0.9850	**0.9815**	0.9770	0.9790	0.9780
Romanian	0.9680	0.9700	0.9690	0.9750	0.9710	**0.9730**
Russian	0.9610	0.9620	**0.9615**	0.9640	0.9570	0.9605
Serbian	0.9750	0.9670	**0.9710**	0.9710	0.9690	0.9700
Slovak	0.9730	0.9690	**0.9710**	0.9710	0.9660	0.9685
Slovenian	0.9470	0.9380	**0.9425**	0.9450	0.9320	0.9385
Spanish	0.9820	0.9810	**0.9815**	0.9810	0.9780	0.9795
Swedish	0.9730	0.9700	0.9715	0.9730	0.9740	**0.9735**
Thai	0.9240	0.9150	**0.9195**	0.9130	0.8980	0.9055
Turkish	0.8780	0.8820	0.8800	0.8910	0.8830	**0.8870**
Ukrainian	0.9270	0.9190	0.9230	0.9260	0.9200	0.9230
Urdu	0.8830	0.8870	0.8850	0.8880	0.8840	**0.8860**
Vietnamese	0.9700	0.9690	**0.9695**	0.9680	0.9640	0.9660

Table 2 presents results from sentence encoding tasks on real-world crisis-related human-labelled social media texts. The evaluated models include SBERT (Sentence Transformer's top-performing model: all-mpnet-base-v2), CT (CrisisTransformers), CT-XLMR-SE, and CT-mBERT-SE. All of these models utilized mean pooling over the token embeddings of an input sequence, considering the attention mask. Across all 18 datasets, our multi-lingual models consistently outperformed SBERT, whose D_{avg} = 0.6374. CT-XLMR-SE with D_{avg} = 0.7255 exhibited a 13.79% improvement over SBERT, and CT-mBERT-SE with D_{avg} = 0.7280 showed a 14.17% improvement. We also present the performance of vanilla XLM-R and mBERT on sentence encoding tasks in Table 3. The results show that sentence embeddings generated by these models lack semantic richness, aligning with findings in the existing literature that transformer-based pre-trained models, out-of-the-box, do not produce semantically meaningful sentence embeddings (Reimers and Gurevych 2019; Lamsal, Read, et al. 2023b).

Table 4 provides a detailed overview of the performance of the student models across 52 languages for sentence matching tasks on the TED2020 test sets. The evaluation metrics include accuracy scores for English to translated language task (en-t) and translated language to English task (t-en) and the average for each language. Out of 52 languages, CT-XLMR-SE performed best in 32 languages on average accuracy, while CT-mBERT-SE excelled in 18 languages. In 2 languages, both models performed equally. The results show variation among models in their performance across languages. Portuguese (Brazil), Spanish, Macedonian, Dutch, Portuguese, German, Serbian, and Slovak exhibit better results with CT-XLMR-SE. However, for languages such as Swedish, Romanian, and Croatian, CT-mBERT-SE surpasses CT-XLMR-SE. For Galician and Ukrainian, both models performed equally. CT-XLMR-SE achieves accuracy above 0.95 across 27 languages, while CT-mBERT-SE achieves this across 28 languages. Seven languages fall within the [0.8, 0.9) accuracy bracket for CT-XLMR-SE, and five languages for CT-mBERT-SE. The only outlier is CT-mBERT-SE, which achieved a score of 0.3055 for Kurdish (Sorani), placing it as the only language below the [0.8, 0.9) bracket. In contrast, CT-XLMR-SE achieved a score of 0.8890 for the same language. Overall, European languages tend to attain higher scores compared to languages from other regions. This trend could be attributed to the availability of more training data for European languages in the parallel datasets considered by this study.

The mean squared error (MSE) loss used in the training is based on the deviation of student's embeddings from teacher's embeddings. Since the student is approximating the embedding space of the teacher, it does not surpass the teacher in tasks involving English texts. The teacher's specialized training on English texts gives it a more nuanced understanding of the language, allowing it to perform well in English tasks. However, the strength of the student model lies in its ability to process multiple languages. Sentence Transformers also offer multi-lingual versions of several of their pre-trained embedding models utilizing the same student-teacher embeddings deviation-based loss. Both our student models surpass Sentence Transformers' top-performing model, *all-mpnet-base-v2*, in English sentence encoding tasks while maintaining multi-lingual capacity. As the second part of the training objective, i.e. $(\mathrm{CT}\ (en_j) - \mathrm{S}(t_j))$, involves aligning the student's non-English embeddings with the teacher's English embeddings, the student produces similar embeddings for translated pairs of sentences, which is evident with results and examples provided in Tables 4, 5 and 6. Table 5 lists an originally posted tweet and its translated (with Google Translate) versions, and Table 6 provides cosine similarities amongst them. Given that our student models replicate the embedding space of CrisisTransformers and surpass Sentence Transformers by > 13%, it is reasonable to expect that they perform well compared to Sentence Transformers' multi-lingual models. However, further study is required to quantify how well these models perform on real-world non-English crisis-related social media texts.

Multilingual embedding models are critical within crisis informatics. Areas where the proposed student models can be applied include semantic search, clustering, and topic modeling. Semantic search powered by multilingual embedding models enables the identification of semantically related content, such as matching help requests with relevant offers (Dutt et al. 2019), thus facilitating efficient allocation of resources during crises. For instance, if a first responding agency needs to filter through a large amount of social media messages, semantic search can quickly help identify messages indicating urgent needs, such as evacuation requests from a building collapse. By comparing the embeddings of search phrases with those of incoming social media streams in multiple languages, relevant messages can be efficiently retrieved and prioritized. For example, a similarity search with embeddings of any sentences in Table 5 with a similarity threshold of 0.9 would return remaining sentences as they are literal translations of each other. Moreover, multilingual embeddings are instrumental in clustering social media messages, offering benefits throughout the disaster management cycle. Clustering helps in organizing similar messages together, aiding in the identification of emerging trends, hotspot areas, or critical needs. This assists decision-makers in allocating resources and planning response strategies effectively. Also, neural topic models (Grootendorst 2022) can use the cross-lingual sentence embeddings to perform topic modeling without language restrictions. These models can uncover underlying themes and topics within crisis-related social media data, providing insights into

evolving situations and public concerns across diverse linguistic contexts.

Table 5. An originally posted crisis-related tweet and its translated versions.

Language	Sentence
	originally posted on Twitter[src]
English	Our state is grappling with severe weather impacts right now, including Melbourne's south east on the back of that terrible storm. I will keep you updated on federal recovery support and please stay safe everyone.
	translated
Chinese (Simplified)	我们的州目前正在努力应对恶劣天气的影响，包括遭受这场可怕风暴影响的墨尔本东南部。我将随时向您通报联邦恢复支持的最新情况，请大家注意安全。
German	Unser Staat ka¨mpft derzeit mit den Auswirkungen schwerer Unwetter, auch im Su¨ dosten Melbournes aufgrund dieses schrecklichen Sturms. Ich werde Sie u¨ ber die Wiederaufbauunterstu¨tzung des Bundes auf dem Laufenden halten und bitte bleiben Sie alle gesund.
Italian	Il nostro Stato e` alle prese con gravi conseguenze meteorologiche in questo momento, compreso il sud-est di Melbourne, a causa di quella terribile tempesta. Vi terro` aggiornati sul supporto federale per la ripresa e per favore state tutti al sicuro.
Japanese	私たちの州は現在、あのひどい嵐の影響でメルボルンの南東部を含む、厳しい気象の影響と闘っています。 連邦政府の復興支援に関する最新情報をお知らせしますので、皆様ご安全にお過ごしください。
Spanish	Nuestro estado esta´ lidiando con impactos clima´ticos severos en este momento, incluido el sureste de Melbourne a ra´ız de esa terrible tormenta. Los mantendre´ informados sobre el apoyo federal a la recuperacio´n y por favor mante´nganse a salvo todos.

[src] https://twitter.com/ClareONeilMP/status/1755717855694954944

Table 6. Cosine similarities amongst the sentences listed in Table 5 (encoded by CT-XLMR-SE).

	Chinese	English	German	Italian	Japanese	Spanish
Chinese	–	0.934	0.931	0.930	0.939	0.937
English	0.934	–	0.943	0.969	0.944	0.970
German	0.931	0.943	–	0.956	0.936	0.947
Italian	0.930	0.969	0.956	–	0.959	0.970
Japanese	0.939	0.944	0.936	0.959	–	0.952
Spanish	0.937	0.970	0.947	0.970	0.952	–

CONCLUSION

In this study, we introduced two multi-lingual sentence encoders (CT-XLMR-SE and CT-mBERT-SE) designed for embedding crisis-related social media texts. Both models were trained as students in student-teacher networks, with CrisisTransformers' (mono-lingual) sentence encoder serving as the common teacher. The training process utilized a large-scale dataset comprising 10 different parallel datasets, totalling over 128 million sentence pairs (e.g., en–es, en–de, en–fr, etc.) across 52 languages. The proposed models underwent evaluation through sentence encoding tasks to assess how well they approximated the teacher's vector space and sentence matching tasks to evaluate their multi-lingual capabilities. Results from both tasks demonstrate that the models generalize well to the languages introduced during training and mimic CrisisTransformers' sentence encoder's embedding space. XLM-R and mBERT, upon which the proposed multi-lingual sentence encoders are based, were originally trained on extensive text corpora containing 100 and 104 languages, respectively. We performed additional training to align these models with CrisisTransformers' sentence encoder's vector space, concentrating on 52 specific languages.

Advancing the area: Our objective is to expand language coverage, including low-resource languages. There is increasing interest in distillation techniques to create small yet effective models from large pre-trained ones. This presents an exciting opportunity to distill knowledge from large models for generating state-of-the-art dense embeddings suitable for real-time processing of crisis-related social media texts. Additionally, we aim to integrate CrisisTransformers with distilled knowledge from large language models to enhance the quality of embeddings. Moreover, the crisis informatics domain lacks translation pairs to train a similar student-teacher architecture, suggesting a potential area for future exploration. Could crisis-specific translation pairs potentially outperform crisis-specific or general-purpose pre-trained models that are subsequently fine-tuned on parallel data from the general domain? This requires further investigation.

ACKNOWLEDGEMENTS

This study is supported by the Melbourne Research Scholarship from the University of Melbourne, Australia. This research was undertaken using the LIEF HPC-GPGPU Facility hosted at the University of Melbourne, which was established with the assistance of LIEF Grant LE170100200.

REFERENCES

Agic, Ž. and Vulic, I. (2019). "JW300: A wide-coverage parallel corpus for low-resource languages". In: Association for Computational Linguistics.

Alam, F., Ofli, F., and Imran, M. (2018). "Crisismmd: Multimodal twitter datasets from natural disasters". In: *ICWSM*. Vol. 12.

Alam, F., Ofli, F., Imran, M., and Aupetit, M. (2018). "A twitter tale of three hurricanes: Harvey, irma, and maria". In: *arXiv preprint arXiv:1805.05144*.

Artetxe, M. and Schwenk, H. (2019). "Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond". In: *Transactions of the Association for Computational Linguistics* 7, pp. 597–610.

Ashktorab, Z., Brown, C., Nandi, M., and Culotta, A. (2014). "Tweedr: Mining twitter to inform disaster response." In: *ISCRAM*, pp. 269–272.

Barrault, L., Bojar, O., Costa-Jussa, M. R., Federmann, C., Fishel, M., Graham, Y., Haddow, B., Huck, M., Koehn, P., Malmasi, S., et al. (2019). "Findings of the 2019 conference on machine translation (WMT19)". In: ACL.

Biddle, R., Joshi, A., Liu, S., Paris, C., and Xu, G. (2020). "Leveraging sentiment distributions to distinguish figurative from literal health reports on Twitter". In: *WWW*, pp. 1217–1227.

Caragea, C., McNeese, N. J., Jaiswal, A. R., Traylor, G., Kim, H.-W., Mitra, P., Wu, D., Tapia, A. H., Giles, C. L., Jansen, B. J., et al. (2011). "Classifying text messages for the Haiti earthquake." In: *ISCRAM*. Citeseer.

Cer, D., Yang, Y., Kong, S.-y., Hua, N., Limtiaco, N., John, R. S., Constant, N., Guajardo-Cespedes, M., Yuan, S., Tar, C., et al. (2018). "Universal sentence encoder". In: *arXiv preprint arXiv:1803.11175*.

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzma´n, F., Grave, E., Ott, M., Zettlemoyer, L., and Stoyanov, V. (2019). "Unsupervised cross-lingual representation learning at scale". In: *arXiv preprint arXiv:1911.02116*.

Conneau, A., Kiela, D., Schwenk, H., Barrault, L., and Bordes, A. (Sept. 2017). "Supervised Learning of Universal Sentence Representations from Natural Language Inference Data". In: *EMNLP*. Copenhagen, Denmark: Association for Computational Linguistics, pp. 670–680.

Conneau, A., Lample, G., Ranzato, M., Denoyer, L., and Je´gou, H. (2017). "Word translation without parallel data". In: *arXiv preprint arXiv:1710.04087*.

Cotfas, L.-A., Delcea, C., Roxin, I., Ioanaˇs¸, C., Gherai, D. S., and Tajariol, F. (2021). "The longest month: analyzing COVID-19 vaccination opinions dynamics from tweets in the month following the first vaccine announcement". In: *IEEE Access* 9, pp. 33203–33223.

Curiskis, S. A., Drake, B., Osborn, T. R., and Kennedy, P. J. (2020). "An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit". In: *Information Processing & Management* 57.2, p. 102034.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). "Bert: Pre-training of deep bidirectional transformers for language understanding". In: *arXiv preprint arXiv:1810.04805*.

Dutt, R., Basu, M., Ghosh, K., and Ghosh, S. (2019). "Utilizing microblogs for assisting post-disaster relief operations via matching resource needs and availabilities". In: *Information Processing & Management* 56.5, pp. 1680–1697.

Gao, T., Yao, X., and Chen, D. (Nov. 2021). "SimCSE: Simple Contrastive Learning of Sentence Embeddings". In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, pp. 6894–6910.

Grootendorst, M. (2022). "BERTopic: Neural topic modeling with a class-based TF-IDF procedure". In: *arXiv preprint arXiv:2203.05794*.

Hayawi, K., Shahriar, S., Serhani, M. A., Taleb, I., and Mathew, S. S. (2022). "ANTi-Vax: a novel Twitter dataset for COVID-19 vaccine misinformation detection". In: *Public health* 203, pp. 23–30.

Hill, F., Cho, K., and Korhonen, A. (June 2016). "Learning Distributed Representations of Sentences from Unlabelled Data". In: *NACL*. San Diego, California: Association for Computational Linguistics, pp. 1367–1377.

Hughes, A. L. and Palen, L. (2009). "Twitter adoption and use in mass convergence and emergency events". In: *International journal of emergency management* 6.3-4, pp. 248–260.

Imran, M., Castillo, C., Diaz, F., and Vieweg, S. (2015). "Processing social media messages in mass emergency: A survey". In: *ACM Computing Surveys* 47.4, pp. 1–38.

Imran, M., Castillo, C., Lucas, J., Meier, P., and Vieweg, S. (2014). "AIDR: Artificial intelligence for disaster response". In: *WWW*, pp. 159–162.

Imran, M., Elbassuoni, S., Castillo, C., Diaz, F., and Meier, P. (2013a). "Extracting information nuggets from disaster-Related messages in social media". In: *ISCRAM*.

Imran, M., Elbassuoni, S., Castillo, C., Diaz, F., and Meier, P. (2013b). "Practical extraction of disaster-relevant information from social media". In: *WWW*, pp. 1021–1024.

Imran, M., Mitra, P., and Castillo, C. (May 2016). "Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages". In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. Portoroˇ, Slovenia: European Language Resources Association (ELRA), pp. 1638–1643.

Karisani, P. and Agichtein, E. (2018). "Did you really just have a heart attack? Towards robust detection of personal health mentions in social media". In: *WWW*, pp. 137–146.

Kiros, R., Zhu, Y., Salakhutdinov, R. R., Zemel, R., Urtasun, R., Torralba, A., and Fidler, S. (2015). "Skip-thought vectors". In: *Advances in neural information processing systems* 28.

Klein, A. Z., Magge, A., O'Connor, K., Flores Amaro, J. I., Weissenbacher, D., and Gonzalez Hernandez, G. (2021). "Toward using Twitter for tracking COVID-19: a natural language processing pipeline and exploratory data set". In: *Journal of medical Internet research* 23.1, e25314.

Koehn, P. (2005). "Europarl: A parallel corpus for statistical machine translation". In: *Proceedings of machine translation summit x: papers*, pp. 79–86.

Lamsal, R., Harwood, A., and Read, M. R. (2022a). "Socially enhanced situation awareness from microblogs using artificial intelligence: A survey". In: *ACM Computing Surveys* 55.4, pp. 1–38.

Lamsal, R., Harwood, A., and Read, M. R. (2022b). "Where did you tweet from? Inferring the origin locations of tweets based on contextual information". In: *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, pp. 3935–3944.

Lamsal, R., Read, M. R., and Karunasekera, S. (2023a). "A Twitter narrative of the COVID-19 pandemic in Australia". In: *ISCRAM*.

Lamsal, R., Read, M. R., and Karunasekera, S. (2023b). "CrisisTransformers: Pre-trained language models and sentence encoders for crisis-related social media texts". In: *arXiv preprint arXiv:2309.05494*.

Li, H., Li, X., Caragea, D., and Caragea, C. (2018). "Comparison of word embeddings and sentence encodings as generalized representations for crisis tweet classification tasks". In: *ISCRAM Asia Pacific*.

Lison, P. and Tiedemann, J. (2016). "Opensubtitles2016: Extracting large parallel corpora from movie and tv subtitles". In.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). "Roberta: A robustly optimized bert pretraining approach". In: *arXiv preprint arXiv:1907.11692*.

Mauriello, M. L., Lincoln, T., Hon, G., Simon, D., Jurafsky, D., and Paredes, P. (2021). "Sad: A stress annotated dataset for recognizing everyday stressors in sms-like conversational systems". In: *Extended abstracts of the 2021 CHI conference on human factors in computing systems*, pp. 1–7.

Olteanu, A., Castillo, C., Diaz, F., and Vieweg, S. (2014). "Crisislex: A lexicon for collecting and filtering microblogged communications in crises". In: *ICWSM*. Vol. 8, pp. 376–385.

Poddar, S., Mondal, M., Misra, J., Ganguly, N., and Ghosh, S. (2022). "Winds of Change: Impact of COVID-19 on Vaccine-related Opinions of Twitter users". In: *ICWSM*. Vol. 16, pp. 782–793.

Pourebrahim, N., Sultana, S., Edwards, J., Gochanour, A., and Mohanty, S. (2019). "Understanding communication dynamics on Twitter during natural disasters: A case study of Hurricane Sandy". In: *International journal of disaster risk reduction* 37, p. 101176.

Purohit, H., Castillo, C., Diaz, F., Sheth, A., and Meier, P. (2014). "Emergency-relief coordination on social media: Automatically matching resource requests and offers". In: *First Monday*.

Reimers, N. and Gurevych, I. (2019). "Sentence-bert: Sentence embeddings using siamese bert-networks". In: *arXiv preprint arXiv:1908.10084*.

Reimers, N. and Gurevych, I. (2020). "Making monolingual sentence embeddings multilingual using knowledge distillation". In: *arXiv preprint arXiv:2004.09813*.

Sarcevic, A., Palen, L., White, J., Starbird, K., Bagdouri, M., and Anderson, K. (2012). "" Beacons of hope" in decentralized coordination: learning from on-the-ground medical twitterers during the 2010 Haiti earthquake". In: *CSCW*, pp. 47–56.

Schwenk, H., Chaudhary, V., Sun, S., Gong, H., and Guzma´n, F. (2019). "Wikimatrix: Mining 135m parallel sentences in 1620 language pairs from wikipedia". In: *arXiv preprint arXiv:1907.05791*.

Simon, T., Goldberg, A., and Adini, B. (2015). "Socializing in emergencies—A review of the use of social media in emergency situations". In: *International journal of information management* 35.5, pp. 609–619.

Starbird, K. and Palen, L. (2010). "Pass it on?: Retweeting in mass emergency". In: *I.*

Stieglitz, S., Mirbabaie, M., Ross, B., and Neuberger, C. (2018). "Social media analytics–Challenges in topic discovery, data collection, and data preparation". In: *International journal of information management* 39, pp. 156–168.

Thomson, R., Ito, N., Suda, H., Lin, F., Liu, Y., Hayasaka, R., Isochi, R., and Wang, Z. (2012). "Trusting tweets: The Fukushima disaster and information source credibility on Twitter." In: *ISCRAM*.

Tiedemann, J. (2012). "Parallel data, tools and interfaces in OPUS." In: *Lrec*. Vol. 2012. Citeseer, pp. 2214–2218.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). "Attention is all you need". In: *Advances in neural information processing systems* 30.

Vieweg, S. (2012). "Situational awareness in mass emergency: A behavioral and linguistic analysis of microblogged communications". PhD thesis. University of Colorado at Boulder.

Vieweg, S., Hughes, A. L., Starbird, K., and Palen, L. (2010). "Microblogging during two natural hazards events: what twitter may contribute to situational awareness". In: *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1079–1088.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., et al. (2020). "Transformers: State-of-the-art natural language processing". In: *EMNLP*, pp. 38–45.

Yang, Y., Cer, D., Ahmad, A., Guo, M., Law, J., Constant, N., Abrego, G. H., Yuan, S., Tar, C., Sung, Y.-H., et al. (2019). "Multilingual universal sentence encoder for semantic retrieval". In: *arXiv preprint arXiv:1907.04307*.

Yang, Y., Yuan, S., Cer, D., Kong, S.-y., Constant, N., Pilar, P., Ge, H., Sung, Y.-H., Strope, B., and Kurzweil, R. (July 2018). "Learning Semantic Textual Similarity from Conversations". In: *Proceedings of the Third Workshop on Representation Learning for NLP*. Melbourne, Australia: Association for Computational Linguistics, pp. 164–174.

Image-text crisis tweet categorization: a caption-based approach

Badreddine Farah*
University of Orléans, INSA-CVL, LIFO, EA 4022, F45067 Orléans, France
badreddine.farah@univ-orleans.fr

Guillaume Cleuziou
University of Orléans, INSA-CVL, LIFO, EA 4022, F45067 Orléans, France
guillaume.cleuziou@univ-orleans.fr

Cécile Gracianne
BRGM, F45060 Orléans, France
c.gracianne@brgm.fr

Adel Hafiane
INSA-CVL, University of Orléans, PRISME, EA 4229, F18022 Bourges, France
adel.hafiane@insa-cvl.fr

Anaïs Halftermeyer
University of Orléans, INSA-CVL, LIFO, EA 4022, F45067 Orléans, France
anais.halftermeyer@univ-orleans.fr

Raphaël Canals
University of Orléans, INSA-CVL, PRISME, EA 4229, F45072 Orléans, France
raphael.canals@uoMappers.niv-orleans.fr

ABSTRACT

The growth of social media usage this last decade has made available a massive and valuable volume of multimedia data. However, the lack of large multimodal annotated datasets, along with the inherent noise and the diversity of multimodal relations in this type of data presents challenges for machine learning methods. Unlike classic multimodal data, social media data comes with a large diversity of relations between image and text making the interaction between the two modalities more difficult. Previous research concentrated on fusion strategies with separate encoders for each modality. This paper introduces CMB (Caption-based Multimodal BERT), a method of classifying crisis-related social media posts by utilizing information from both images and texts. CMB translates the image modality into a text-compatible space, facilitating intermodal interaction. Furthermore, CMB presents training opportunities to enhance the model's robustness to missing modalities. Experimental results show that CMB is competitive with well-established, costly, and manually crafted multimodal models.

Keywords

Deep Learning, multimodal data, text/image fusion, Crisis data

INTRODUCTION

Proliferation of social media and advancements in internet technology have significantly changed the nature of online content, especially on platforms such as X (previously Twitter[1]), where user-generated data has immense value to a wide range of domains, including public health, economics, and politics (Wu & Mebane Jr, 2022). In times of crisis (natural disasters), this data can be used to gain situational awareness and assist humanitarian organizations in their preparedness, mitigation, response, and recovery efforts. However, the large volume of data present on social media requires the development and implementation of automated systems to effectively categorize and identify informative posts by analyzing information present across modalities.

Many works in this scope focused on processing unimodal data including images (Nguyen et al., 2017) or texts (Kumar et al., 2011). To leverage both modalities, many papers have proposed general-purpose self-supervised

*corresponding author

1We will be using Twitter and tweets as a reference for X and X posts

vision-language models (Y.-C. Chen et al., 2020; Kim et al., 2021; Tan & Bansal, 2019). Nevertheless, these large pre-trained (multimodal) models are primarily designed for homogeneous data, such as images and their captions. They therefore do not take into account the heterogeneous informational content that may exist between an image and its associated text, as is the case in social media (Vempala & Preo¸tiuc-Pietro, 2019), making transfer learning suboptimal (Liang et al., 2022). In this context of multimodal social media tasks, prior works have concentrated their efforts on fusion strategies. The prevailing practice involves using two unimodal encoders, one for each modality, followed by fusing the unimodal embeddings. Many fusion strategies were proposed (Ofli et al., 2020) for instance used the concatenation of the two embeddings, while (Abavisani et al., 2020) introduced a cross-attention mechanism to achieve a more effective fusion. In contrast, (Liang et al., 2022) explored early interaction strategies to enable interactions on low-level features.

However, the lack of large annotated datasets coupled with the variety of the inter-modality relations between image and text presents additional challenges for learning task-oriented fusion. Moreover, previous studies (Fan et al., 2023; Wang et al., 2020) have shown that different modalities tend to overfit and converge at different rates when using two encoders. This indicates that optimizing the same objective for various modalities results in inconsistent learning efficiency. In this paper, we investigate the effectiveness of captions in vision-language tasks for crisis tweets. It is hypothesized that the embedding integration of both texts and images into a common semantic space and the use of a single trainable encoder may thus facilitate the challenging learning problem. For this, we propose to use a caption-based method, which consists of translating a modality (e.g. image) towards another modality (e.g. text). We, therefore, present the CMB method ***C****aption-based* ***M****ultimodal* ***B****ERT*, that leverages captioning models to translate image into text and concatenate resulting captions with the tweet's text before inputting the combined data into a text encoder.

To thoroughly evaluate this approach, we conduct experiments with a variety of captioning models that have been pre-trained on different datasets. We also propose a training strategy that allows to use our method in unimodal settings without significant performance loss. Our experiments focus on crisis context, in particular on the well-studied multimodal CrisisMMD dataset (Alam et al., 2018). The examples in Figures 4 and 5 illustrate the challenges of this dataset in terms of the relationship between modalities (complementarity, heterogeneity or partial redundancy (Vempala & Preo¸tiuc-Pietro, 2019)). Our contributions are as follows:

- We present CMB, a method to homogenize the input space of image/text data to enable better information processing by BERT-like architecture.
- We compare three captioning models and show that this method can achieve highly competitive results for crisis-related tweets.
- We propose a training strategy to use our models in unimodal settings without significant loss of performance.
- Lastly, we present a qualitative analysis of the success and failure of our method.

RELATED WORKS

IMAGE-TEXT MULTIMODAL LEARNING

The image-text machine learning research area encompasses a variety of application domains, such as image captioning (Cho et al., 2022; Mokady et al., 2021), visual question answering (Antol et al., 2015) and visual reasoning (Zellers et al., 2019). Recently, several pre-trained models have been developed to address these tasks : LXMERT (Tan & Bansal, 2019) employs two unimodal streams and a cross-attention encoder to combine information from multiple modalities, while UNITER (Y.-C. Chen et al., 2020) uses a single stream transformer to enable inter-modality attention from the early layers of the model. In contrast, VILT (Kim et al., 2021) allows interaction between text and image at the fine-grained feature level without requiring prior convolution or region detection.

Multimodal learning has been applied to a range of social media projects, including emergency response (Abavisani et al., 2020; Liang et al., 2022; Long & McCreadie, 2022), emotional recognition (Yang et al., 2020), fake news detection (Tuan & Minh, 2021) and political science (Wu & Mebane Jr, 2022). There are two main approaches to multimodal social media post classification: early fusion and late fusion. In early fusion, the interaction between modalities is allowed from the early layers of the model, as shown by the Multimodal Information Injection Plug-in units introduced in (Liang et al., 2022), which preserve the strong intra-modal processing ability of large pre-trained unimodal models. In late fusion, separate encoders are used for each modality, such as DenseNet (Huang et al., 2017) or ResNet (He et al., 2016) for images and BERT (Devlin et al., 2018) or RoBERTa (Liu et al., 2019) for texts.

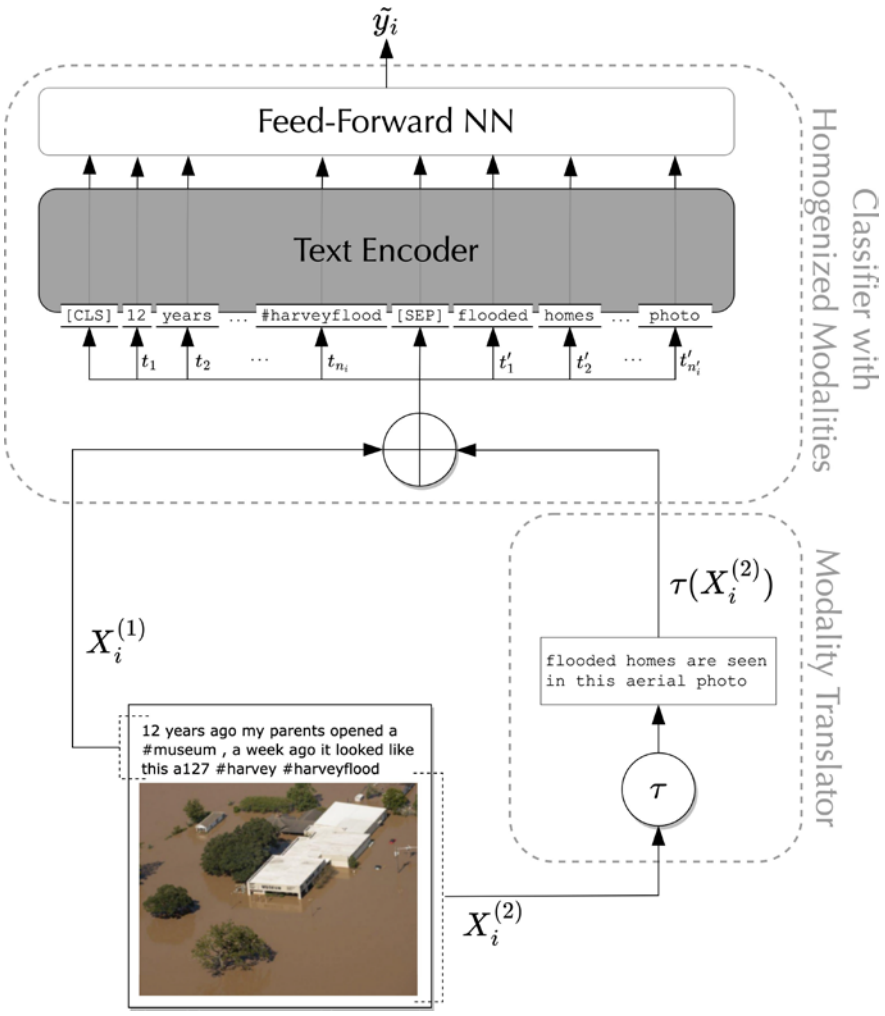


Figure 1. Overview of the CMB method structured in two main parts: a non-learnable process represented by the modality translator (image captioning model) and the homogenized modalities classifier as trainable module.

Interaction between modalities occurs at the high feature level through techniques such as feature concatenation (Long & McCreadie, 2022) or cross-attention (Abavisani et al., 2020).

In this study, however, we explore the effectiveness of image captions, employing pre-trained captioning model to capture the image information. The acquired information is then integrated into a text embedding space to facilitate fusion.

MACHINE LEARNING FOR EMERGENCY RESPONSE

Using social media information in emergency response is increasingly considered essential for the management of natural disaster crises. Many prior works have focused on unimodal signals: (Kumar et al., 2011) proposes a system to analyze and monitor (text) tweets to help humanitarian organizations better manage crises, (Nguyen et al., 2017) uses images posted on Twitter to determine the level of damage caused by disasters (Little, mild or severe). In a multimodal setting, (Abavisani et al., 2020) applies late fusion with two unimodal encoders to filter informative image-text tweets, while (Liang et al., 2022) employs early fusion to enhance low-level inter-modal feature interactions. Unlike previous approaches, we argue that the inherent complexity within social media data can be challenging due to the verity of relations between modalities, (Vempala & Preo¸tiuc-Pietro, 2019). To overcome this challenge (Sosea et al., 2021) introduced DisRel, a dataset for learning to distinguish the relationships between images and texts according to their nature (*complementary*, *similar* or *unrelated*). Their work demonstrated that incorporating such tasks can enhance the performance of models trained on CrisisMMD.

In this work, we propose an alternative method leveraging captioning techniques for image-to-text translation, in order to map the image representation space into text facilitating fusion within a shared representation space.

IMAGE CAPTIONING FOR FUSION

In previous works, captions have been used to represent images for multimodal interaction, such as Multimodal Named Entity Recognition (MNER) (S. Chen et al., 2020). It has been argued that models trained for semantic understanding, like image captioning models, may provide better image representations. The use of captions as image representations allows for the incorporation of textual descriptions of the image content, providing additional context and useful information to understand and interpret the image. Moreover, (Wu & Mebane Jr, 2022) uses captions as a translated version of image in addition to the tweet's image and text to improve performance in classifying multimodal tweets. Contrary to (Wu & Mebane Jr, 2022), the method CMB proposed in this paper uses - through captions - modality translation as a substitute and not as a complement to the image to improve inter-modality interactions while keeping the modal simple to train and use in real-world scenarios (with limited resources). Moreover, we argue that the use of one encoder helps to overcome the problem of different convergence rates of each modality encoder and lead to more efficient learning.

METHODOLOGY

PROBLEM STATEMENT

Let D be a multimodal dataset with N instances, denoted by $D = \{(X_i, y_i)\}_{i=1}^{N}$. Each $X_i \in X$ represents a multimodal data point with M modalities, represented as $X_i = \{X_i^{(p)}\}_{p=1}^{M}$. Associated with each instance is a label $y_i \in C$, where C is the set of possible labels. The multimodal classification task aims at learning a classifier $h : X \rightarrow C$ that combines information from different modalities to make accurate predictions. In the following, we consider only bimodal instances ($M = 2$) where the two modalities are different in nature, typically textual and visual modalities. Thus, $X_i^{(1)} = (t_1, \ldots, t_{n_i})$ is a sequence of tokens while $X_i^{(2)} \in R^{C \times H \times W}$ is an image (where C, H, and W are the number of channels, the height, and the width of the image respectively.

MODEL OVERVIEW

Since modalities of different natures (typically text and image) are hard to combine in multimodal fusion strategies, we propose a two-step categorization approach that leverages modality translation to homogenize the inputs (Figure 1). The first step, *Modality Translator* consists in translating one modality into a space compatible with the space of the other modality using a translator function $\tau(.)$. In our particular context of text-image tweets processing, the image modality $X_i^{(2)}$ is translated into a text (sequence of tokens) by means of an image captioning process $\tau(X_i^{(2)}) = (t'_1, \ldots, t'_{n'_i})$. The second step *Classifier with Homogenized Modalities* takes as input the concatenation of the homogenized modalities $X_i^{(1)} \oplus \tau(X_i^{(2)})$ to train both a text encoder and the final classification layer. Let us notice that the approach allows to process both multimodal ($X_i^{(1)} \oplus \tau(X_i^{(2)})$) or unimodal data ($X_i^{(1)}$ only or $\tau(X_i^{(2)})$ only).

MODALITY TRANSLATORS

In our multimodal approach, the captioning model is viewed as a modality translator. For our investigation, we use three publicly available captioning models, which vary depending on the encoders and the training data they use.

- **CLIP Prefix for Image Captioning** (Mokady et al., 2021). A captioning model using the CLIP (Radford et al., 2021) (Contrastive Language-Image Pre-training) multimodal encoder, which is trained on a large dataset sourced from the Web. It extracts image features using the CLIP encoder and employs them as a prefix for text generation by a language model. Trained with either the MSCOCO dataset (Lin et al., 2014) or the Conceptual Caption dataset (Sharma et al., 2018), the authors argue that the rich semantic features of the CLIP encoder, coupled with a pre-trained language model (GPT2 (Radford et al., 2019)), provide a comprehensive understanding of both visual and textual data, resulting in high quality captions. In the following, these captioning models are referred to as *CLIP_cc* (trained with Conceptual Captions) and *CLIP_coco* (trained with MSCOCO).
- **Fine-grained Image Captioning with CLIP Reward**. In (Cho et al., 2022) the authors use CLIP (Radford et al., 2021) in two ways, first as an image encoder, and then CLIP text and CLIP image are used to compute a multimodal similarity that then serves as a reward function for each generated caption to obtain descriptive and distinctive captions. This captioning model is referred to as *CLIP_reward* in the following.

Table 1. Training and evaluation settings.

Train \ Eval	*Text*	*Image*	*Caption*	*Text + Image*	*Text ⊕ Caption*
Text	T_T	-	-	-	-
Image	-	I_I	-	-	-
Caption	-	-	C_C	-	-
Text + Image	-	-	-	TI_TI	-
Text ⊕ Caption	TC_T	-	TC_C	-	TC_TC

- **Transformer based Captioning**.(Vaswani et al., 2017) uses the Transformer architecture and first process the image using Mask-RCNN (He et al., 2017), the features are then fed to the Transformer in order to generate the caption. This captioning model is referred to as *Trans cap* in the rest of the article.

The CMB approach we propose in this paper allows to independently integrate any of the four captioning models presented above. In the following experiments, we will focus on measuring the contribution of the captions generated by each model in our proposed multimodal framework.

MULTIMODAL CLASSIFIER WITH HOMOGENIZED MODALITIES

Transformer based models (Devlin et al., 2018; Liu et al., 2019; Radford et al., 2019; Vaswani et al., 2017) has achieved state of the art result in large set of natural language processing tasks, by leveraging various self-supervised pre-training techniques. Therefore, without loss of generality, we use BERT (Devlin et al., 2018) as text encoder to extract embeddings from the input $(X_i^{(1)} \oplus \tau(X_i^{(2)}))$, which are then used to perform the classification task using a feed forward neural network layer.

EXPERIMENTAL SETUP

DATASET AND TASKS

The **CrisisMMD** (Alam et al., 2018) benchmark aims to identify crisis events that require emergency response using social media posts as a basis. The dataset for this benchmark consists of tweets (image-text pairs) obtained through searching for specific hashtags on Twitter and labeled for three tasks Informativeness, Humanitarian, and Damage severity assessment. In this work, we only consider the first two multimodal tasks:

- **Informativeness**: a binary classification task identifying whether or not a given tweet (text or image) is informative for humanitarian aid purposes, i.e., useful for providing assistance to people in need.
- **Humanitarian**: Identifying whether a given tweet (text or image) belongs to one of the following seven categories : *infrastructure and utility damage*; *vehicle damage*; *rescue, volunteering, or donation efforts*; *injured or dead people*; *missing or found people*; *other relevant information*; *not humanitarian related*.

This dataset provides a total of 16,058 tweet texts and 18,082 tweet images labelled for these two tasks, resulting into 8,079 multimodal instances for the Humanitarian task and 12,168 instances for the Informativeness task. Unfortunately, the CrisisMMD dataset does not offer one label for each *tweet*, but one label for each *modality* of a tweet. Therefore, it is not directly suited for the (multimodal) classification task we are considering in this work. Thus, we adopt the practice proposed in (Ofli et al., 2020) by using only instances with similar labels on the two modalities (text and image). Although this process excludes a non-negligible number of tweets, it is nevertheless based on the fairly reasonable assumption that a label that matches on both modalities can be used as a label for the entire tweet.

BASELINES

We compare our approach to both uni- and multimodal baselines in the field of image-text classification. Recently, the trend is to use multiple architectures in this research area. To provide a comprehensive evaluation of our approach, we implemented two popular multimodal architectures: the feature concatenation (Long & McCreadie, 2022) and the cross-attention (Abavisani et al., 2020) methods. Additionally, we included the unimodal methods BERT (Devlin

Table 2. Comparisons on CrisisMMD in terms of classification accuracy, Macro F1-score and weighted F1-score.

Models		Informativeness			Humanitarian		
		Acc	F1-m	F1-w	Acc	F1-m	F1-w
Unimodal	BERT	0.8691±0.0031	0.8465±0.0041	0.8667±0.0033	0.8183±0.0068	0.6964±0.0271	0.8179±0.0069
	DenseNet	0.8393±0.0028	0.8186±0.0022	0.8396±0.0024	0.7608±0.0178	0.5366±0.0495	0.7416±0.0272
	CLIP	0.8594±0.0281	0.8407±0.0283	0.8591±0.0265	0.8136±0.0224	0.6105±0.0404	0.8005±0.0325
Multimodal	Features-concat	0.9014±0.0020	0.8870±0.0031	0.9008±0.0023	0.8561±0.0102	0.6772±0.0070	0.8533±0.0098
	Cross-Attention	0.9022±0.0037	0.8885±0.0043	0.9019±0.0037	0.8549±0.0060	0.6718±0.0059	0.8520±0.0058
	CLIP-Concat	0.9020±0.0045	0.8871±0.0054	0.9012±0.0046	0.8608±0.0080	0.7778±0.0146	0.8589±0.0077
CMB	CMB-Trans_cap	0.8920±0.0039	0.8749±0.0045	0.8907±0.0039	0.8527±0.0065	0.7536±0.0274	0.8520±0.0070
	CMB-CLIP-reward	0.9125±0.0033	0.8987±0.0040	0.9115±0.0034	0.8765±0.0080	0.7631±0.0266	0.8755±0.0079
	CMB-CLIP-coco	0.9098±0.0024	0.8954±0.0029	0.9087±0.0025	0.8681±0.0064	0.7156±0.0307	0.8660±0.0066
	CMB-CLIP-cc	**0.9203±0.0036**	**0.9081±0.0043**	**0.9196±0.0037**	**0.8827±0.0033**	**0.7810±0.0176**	**0.8826±0.0032**

et al., 2018) (text only) and DenseNet (Huang et al., 2017) (image only) for comparison. Furthermore, we conduct a comparison between our approach and two other methods: CLIP (Radford et al., 2021) and CLIP-Concat. In the latter, we concatenate features from CLIP image encoder and BERT to accomplish the multimodal classification task. We aim to evaluate the effectiveness of CMB against these well-established methods to illustrate its usefulness in the crisis domain.

IMPLEMENTATION DETAILS

In this work we consider the four modality translators described above. Specifically, for each captioning model, we produced five captions for every image using the authors' provided demo for CLIP-based captioning models (CLIP_cc, CLIP coco and CLIP reward) and the demo from a GitHub repository2 for the Trans_cap model. We keep the most similar caption - from the five generated by each model - to the image in terms of clip similarity score (Cho et al., 2022).

We used the BERT base model from (Wolf et al., 2020) as a text encoder with token type embeddings different for captions ($\tau(X_i^{(2)})$) and texts ($X_i^{(1)}$). To classify, we computed the mean pooling on embeddings provided by BERT and used a classification layer consisting of a ReLU activation function and a dropout layer between two feed-forward layers. The models were trained on a Nvidia Tesla V100 GPU for 3 to 5 epochs with $batch\ size = 128$ and $learning\ rate = 5e^{-5}$. We made our code available on GitHub3.

UNIMODAL AND MULTIMODAL SETTINGS

To evaluate the robustness of our methodology across diverse scenarios, we conduct an array of experiments involving the training and evaluation of various iterations of our model, each employing distinct modalities such as Text, Image, or Caption. Table 1 uses specific notations to delineate these different experimental setups.

The first three rows of the table focus on unimodal approaches (T_T, I_I, and C_C settings), wherein a single modality is exclusively used for both training and evaluation. For text categorization (T_T, C_C), we employ BERT (Devlin et al., 2018), while for image categorization (I_I), DenseNet (Huang et al., 2017) is utilized.

The fourth row shows multimodal approaches (TI_TI setting), involving the simultaneous use of Text and Image during both training and evaluation. This incorporates a basic feature concatenation strategy, along with the previously mentioned cross-attention architecture (Abavisani et al., 2020) which uses attention mechanisms to merge representations from the two unimodal architectures (BERT and DenseNet).

Finally, the last line corresponds to our multimodal method based on modality translation, whose architecture is generic enough to adapt to multiple settings. We investigate its ability to learn efficient hybrid models for categorizing both multimodal and unimodal data by training it on both text and caption and then evaluate it on different data configurations: Text only (TC_T), Caption only (TC_C) and Text $\oplus$ Caption (TC_TC).

We utilize the same dataset across various configurations. In the TC configurations, both text and caption are used as input during training or testing phases. In the T configuration, only text is used as input to the model, with captions being ignored. Conversely, in the C configuration, only captions are used as input, while text is disregarded. This data split remains consistent across all configurations, encompassing training, validation, and testing phases.

To expand our experiments, we propose a mixed training strategy in which the model is trained on a mixture of data that contains text-only instances ($X_i^{(1)}$) and on instances with Text $\oplus$ Caption ($X_i^{(1)} \oplus \tau(X_i^{(2)})$). The details of our experiments are presented in the following sections.

2https://github.com/ruotianluo/self-critical.pytorch.
3anonymous

Table 3. Comparison with MARMOT (Wu & Mebane Jr, 2022), a multimodal model integrating caption, text, and image modalities. Two captioning models, namely *Trans_cap* and *CLIP_cc*, are employed. The parameter count (in millions) is reported, excluding parameters associated with captioning models.

	Informativeness			
	Acc	F1-m	F1-w	Parameter count
MARMOT Trans_cap	0.9133±0.0018	0.9012±0.0021	0.9131±0.0018	309
CMB-Trans_cap	0.8920±0.0039	0.8749±0.0045	0.8907±0.0039	112
MARMOT CLIP_cc	0.9220±0.0022	0.9105±0.0024	0.9215±0.0021	309
CMB-CLIP_cc	0.9203±0.0036	0.9081±0.0043	0.9196±0.0037	112

EXPERIMENTAL RESULTS

This section presents the results of our experiments on the proposed method using different captioning models compared with unimodal and multimodal baselines. Additionally, we evaluate the robustness of CMB on unimodal data settings.

CAPTIONING MODELS COMPARISON

Table 2 displays the results of our experiments on the two multimodal tasks of the CrisisMMD dataset (Informativeness and Humanitarian). The first two rows report unimodal classifiers results, while the next two rows present the performances of multimodal baselines. The last four rows show the performances of our approach using each of the four captioning models described before. Firstly, the results obtained confirm that multimodal methods outperform unimodal ones. Additionally, our method is very competitive with other multimodal baselines, regardless of the modality translator used.

Specially, the performance gap between the different captioning models can be significant, as demonstrated by the 3% performance gap on the informativeness task between the method based on Trans_cap and the best-performing method. Importantly, the CLIP_cc captioning model allows the CMB method to outperform the proposed strong baseline of 1-3% in terms of weighted F1-score on both CrisisMMD tasks. Moreover, employing the identical data splits, our experiments yield competitive results with state-of-the-art (Liang et al., 2022) findings as well as results reported on the MVSA dataset. In the tasks of informativeness and humanitarian, we scored 91.96 and 94.84 W-F1, respectively, while their scores were 91.3 and 93.6.

Furthermore, we can observe the gap between the CLIP_cc and CLIP_coco models despite using the same architecture: we hypothesize that the Conceptual Caption dataset (Sharma et al., 2018) has more coverage and similarity with Twitter images, containing figurative representations with different levels of iconicity such as photographs, drawings, charts, or even non-figurative pieces of digital art for instance. Finally, the clear gap between the CLIP-based models and the simple Transformer model confirms the findings of (Bielawski et al., 2022), indicating that CLIP generalizes better on human-centric tasks (defined by the authors as involving knowledge of cultural, social, aesthetic and/or affective components of the world). In our case, this fits particularly our need on social media, where humans post for other humans.

In order to expand our investigations, we compared our approach with MARMOT (Wu & Mebane Jr, 2022) : a model incorporating images, captions, and textual elements for tweets classification. Employing two distinct captioning models, namely Trans_cap and CLIP_cc, the experiments outcomes are shown in Table 3. Despite the diminished parameter count (approximately one-third) in our method relative to MARMOT, it exhibited competitive performance in both scenarios. Noteworthy disparities were observed, particularly in the Trans_cap scenario, whereas the performance differentials were closely aligned when utilizing CLIP_cc captions. This underlines the influence of employing high-quality captions in this configuration. In the continuation of this paper, we use the best captioning model, namely CLIP cc.

UNIMODAL AND MULTIMODAL RESULTS

As discussed previously, the CMB method contains two parts : a modality translator and a trainable module (text encoder + classifier). The input of the trainable part is exclusively text-based, enabling the use of both unimodal ($X_i^{(1)}$ or $\tau(X_i^{(2)})$) and multimodal ($X_i^{(1)} \oplus \tau(X_i^{(2)})$) data as input. This section investigates the unimodal and multimodal capabilities of the model by conducting an experiment with three main configurations, TC_T, TC_C and TC_TC. Results of the experiments are reported in Figure 2. Performances of CMB (for the different

4We used significance level = 0.05

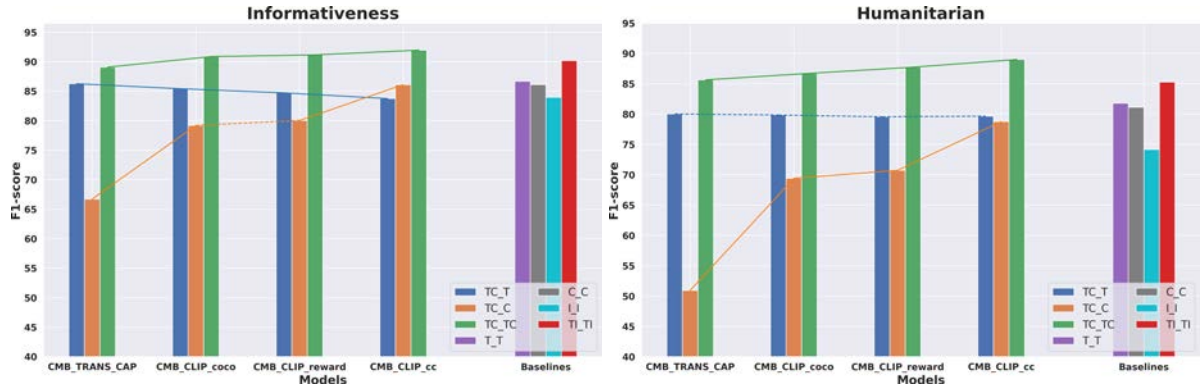


Figure 2. Comparison of modality translators on Unimodal and Multimodal settings (see Table 1). Y-axis shows mean Weighted F1-score results over 10 runs. The X-axis represents the different modality translators used in this experiment. Dotted lines are used for non-significant differences.4

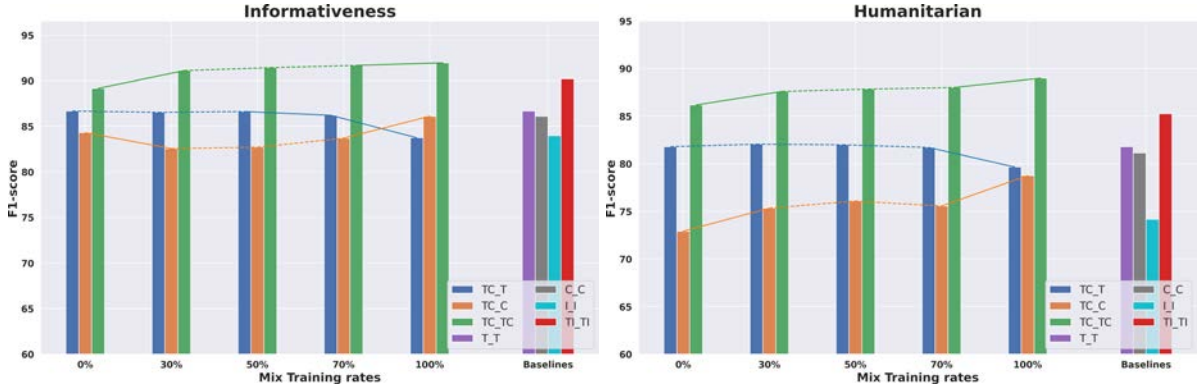


Figure 3. Mix training results. In the X-axis, $X\%$ represents the proportion of multimodal instances ($X_i^{(1)} \oplus \tau(X_i^{(2)})$) in training data, the remaining $100 - X\%$ data are text only ($X_i^{(1)}$). Dotted lines are used for non-significant differences.

captioning models) is compared across the three main configurations as well as to BERT (Devlin et al., 2018) (as T_T configuration), CMB_CLIP_cc (C_C), DenseNet (Huang et al., 2017) (I_I) and the features concatenation model (TI_TI). As mentioned in the previous section, the results clearly illustrate the improved performance when using the two modalities sources for classification. Additionally, these experiments reveal a relationship between the captions-only (TC_C) and multimodal (TC_TC) performances, indicating that the model performance depends on the caption quality. The figures also show that captions-only models can outperform the (I_I) DenseNet model. However, the model in the TC_T setting exhibits poor performance compared to BERT (in T_T configuration) with a 2-3% loss. These results suggest a form of *caption dependency* of the models. To address this issue, we discuss a mix training method in the next section.

MIX TRAINING RESULTS

Twitter posts can be either unimodal or multimodal, but in practice, keywords are used as filters to retrieve tweets with the Twitter API, making some tweets containing only text (without image). Rather than using two different models depending on the nature of each tweet (unimodal or multimodal), we suggest that the CMB model can be used to process both unimodal and multimodal tweets. This hybridization requires an appropriate training strategy, called the *mix training* approach. In this study, we propose an experiment to test our method in TC_T and T_T settings. The goal of this mix training approach is to train a hybrid model that can handle both multimodal (*text+caption*) and unimodal (*text-only*) tweets. To this end, we consider training samples with various *text-only*/*text+caption* balances and observe the ability of the derived models to classify *text-only*, or *text+caption* tweets. In this experiment, the captioning model used is CLIP_cc. The proportions of *text+caption* tweets in the training sample is increased from 0% to 100%, with the remaining examples being *text-only* tweets. The results of the mix-training strategy, as presented in Figure 3, provide interesting insights into the performance of the hybrid model in different scenarios.



Figure 4. Examples of caption using success cases when baseline BERT and Cross-attention failed

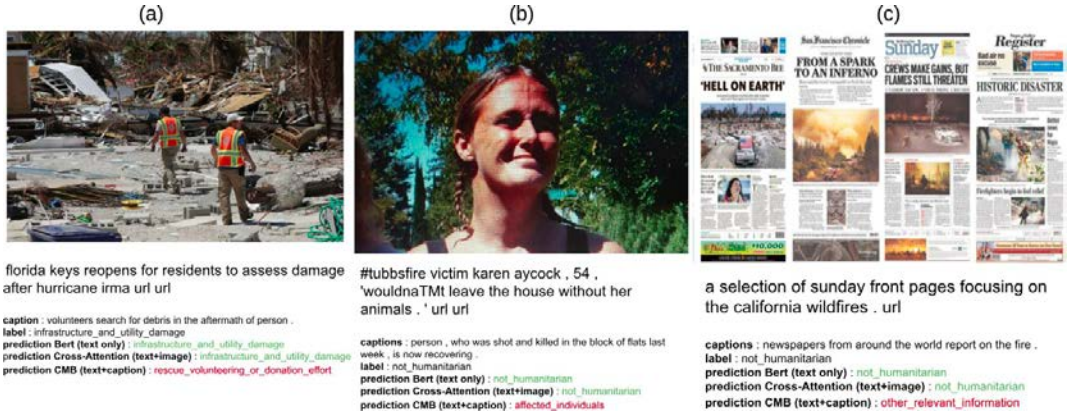


Figure 5. Examples of caption leading to failure cases when baseline BERT and Cross-attention succeed.

When evaluated on text-only tweets (TC_T), the model's performance (F1-score) experiences a slight decrease of approximately 2-3% as the proportion of text+caption tweets in the training sample increases from 50% to 100%. This indicates that the introduction of captions has affected the performance of the model on text-only settings, specially when having more than 70% *text+caption* examples. Additionally, it's evident that the performance in caption-only evaluation (indicated by the orange bar in TC_C) improves as the proportion of *text+caption* examples increases. However, there is one unexplained anomaly observed during the informativeness experiment when transitioning from 0% to 30% of *text+caption* examples.

On the other hand, in the multimodal setting (TC_TC), the same change in the training sample, with an increase in the proportion of text+caption tweets from 50% to 100%, results in an improvement of about 0.5-1% in the model's performance. This suggests that the hybrid model benefits from exposure to a higher proportion of multimodal data during training, leading to better performance in the multimodal tweet classification task. The proposed hybrid model demonstrates its competitiveness with BERT (T_T) when trained on a sample with 50% text+caption tweets, achieving an impressive F1-score of 81.98%. Despite a small loss of performance in the TC_TC setting compared to the text-only scenario, the hybrid model remains clearly superior to the TI_TI baseline, showcasing its capability to effectively handle both unimodal and multimodal tweets. Furthermore, the analysis reveals that the best compromise on caption rates, is achieved at 50%. This suggests that a balanced mixture of text-only and text+caption tweets in the training sample is optimal for achieving peak performance in the hybrid model. Finally, this experiment indicates that the mix training strategy makes the learned models more robust and usable for classifying both unimodal and multimodal tweets in real time.

QUALITATIVE ANALYSIS

Finally, we provide several examples to illustrate the success and failure cases of our proposed method. Figure 4 shows three examples in which the cross-attention baseline and BERT failed, but our method made the correct prediction on the Humanitarian task. In particular, Figure 4(a) is an example of the success of using a caption, as the synthetic image is difficult to be processed by a conventional image model pre-trained on ImageNet (Deng et al., 2009). In figure 4(b) the successful generation of the words "blocks the road" may have contributed to the prediction of the correct label, highlighting the model's ability to capture discriminative information. Finally, Figure 4 (c) is a multimodal example in which the text alone fails to predict the label; in this case, the caption captures information that is merged with the text to achieve the right prediction.

In Figure 5, we present three examples where the caption model produced inaccurate results. The image in Figure 5(a) illustrates the potential of captions to be misleading and result in incorrect predictions; the model captured some information from the image in a poorly structured sentence, resulting in an incorrect prediction. The image in Figure 5(b) demonstrates a failure due to an inaccurate caption generated by the captioning model, highlighting the sensitivity of our approach to the modality translator. Finally, the image in Figure 5(c) shows a situation in which both the label and the prediction fall into a gray area, where either could be considered a valid label. These examples provide insight into the successes and failures of our proposed method, specifying potential areas for improvement.

CONCLUSION

In conclusion, our multimodal classification method based on a modality translation process was found to be competitive with strong baselines and state-of-the-art results on the CrisisMMD dataset. Our analysis confirms that multimodal approaches generally outperform unimodal ones and that the performance gap between different captioning models can be significant. We also found that the caption quality is important for the overall performance of the model, and we proposed a mixed training method to create a hybrid model that can handle both unimodal and multimodal sources. However, our method has several limitations that should be addressed in future work, including the sensitivity on caption quality and the focus on the crisis domain. Overall, CMB is a promising approach for multimodal classification and highlights the importance of considering multiple modalities in tasks involving social media data.

ACKNOWLEDGEMENT

This work was supported by the National Research Agency under the program IA.iO (ANR-20-THIA-0017-01). The computation was performed using Leto resources from CASciModOT federation.

REFERENCES

Abavisani, M., Wu, L., Hu, S., Tetreault, J., & Jaimes, A. (2020). Multimodal categorization of crisis events in social media. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14679–14689.

Alam, F., Ofli, F., & Imran, M. (2018). Crisismmd: Multimodal twitter datasets from natural disasters. *Twelfth international AAAI conference on web and social media*.

Antol, S., Agrawal, A., Lu, J., Mitchell, M., Batra, D., Zitnick, C. L., & Parikh, D. (2015). Vqa: Visual question answering. *Proceedings of the IEEE international conf. on computer vision*, 2425–2433.

Bielawski, R., Devillers, B., Van De Cruys, T., & Vanrullen, R. (2022). When does CLIP generalize better than unimodal models? when judging human-centric concepts. *Proceedings of the 7th Workshop on Representation Learning for NLP*, 29–38. https://doi.org/10.18653/v1/2022.repl4nlp-1.4

Chen, S., Aguilar, G., Neves, L., & Solorio, T. (2020). Can images help recognize entities? a study of the role of images for multimodal ner. *arXiv preprint arXiv:2010.12712*.

Chen, Y.-C., Li, L., Yu, L., El Kholy, A., Ahmed, F., Gan, Z., Cheng, Y., & Liu, J. (2020). Uniter: Universal image-text representation learning. *Computer Vision–ECCV 2020: 16th European Conf., Glasgow, UK, August 23–28, 2020, Proceedings, Part XXX*, 104–120.

Cho, J., Yoon, S., Kale, A., Dernoncourt, F., Bui, T., & Bansal, M. (2022). Fine-grained image captioning with clip reward. *arXiv preprint arXiv:2205.13115*.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Fan, Y., Xu, W., Wang, H., Wang, J., & Guo, S. (2023). Pmr: Prototypical modal rebalance for multimodal learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 20029–20038.

He, K., Gkioxari, G., Dolla´r, P., & Girshick, R. (2017). Mask r-cnn. *Proc. of the IEEE int. conf. on computer vision*, 2961–2969.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4700–4708.

Kim, W., Son, B., & Kim, I. (2021). Vilt: Vision-and-language transformer without convolution or region supervision. *International Conference on Machine Learning*, 5583–5594.

Kumar, S., Barbier, G., Abbasi, M. A., & Liu, H. (2011). Tweettracker: An analysis tool for humanitarian and disaster relief. *Proceedings of the International AAAI Conference on Web and Social Media*.

Liang, T., Lin, G., Wan, M., Li, T., Ma, G., & Lv, F. (2022). Expanding large pre-trained unimodal models with multimodal information injection for image-text multimodal classification. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15492–15501.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dolla´r, P., & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. *European conf. on computer vision*, 740–755.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). Roberta: A robustly optimized BERT pretraining approach. *CoRR, abs/1907.11692*.

Long, Z., & McCreadie, R. (2022). Is multi-modal data key for crisis content categorization on social media? *ISCRAM 2022 Conference Proceedings 226 19th International Conference on Information Systems for Crisis Response and Management*, 1068–1080.

Mokady, R., Hertz, A., & Bermano, A. H. (2021). Clipcap: Clip prefix for image captioning. *arXiv preprint arXiv:2111.09734*.

Nguyen, D. T., Ofli, F., Imran, M., & Mitra, P. (2017). Damage assessment from social media imagery data during disasters. *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 569–576. https://doi.org/10.1145/3110025.3110109

Ofli, F., Alam, F., & Imran, M. (2020). Analysis of social media data using multimodal deep learning for disaster response. *arXiv preprint arXiv:2004.11838*.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., et al. (2021). Learning transferable visual models from natural language supervision. *International Conference on Machine Learning*, 8748–8763.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, *1*(8), 9.

Sharma, P., Ding, N., Goodman, S., & Soricut, R. (2018). Conceptual captions: A cleaned, hypernymed, image alt-text dataset for automatic image captioning. *Proc. of the 56th Annual Meeting of the Association for Computational Linguistics*, 2556–2565.

Sosea, T., Sirbu, I., Caragea, C., Caragea, D., & Rebedea, T. (2021). Using the image-text relationship to improve multimodal disaster tweet classification. *The 18th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2021)*.

Tan, H., & Bansal, M. (2019). Lxmert: Learning cross-modality encoder representations from transformers. *arXiv preprint arXiv:1908.07490*.

Tuan, N. M. D., & Minh, P. Q. N. (2021). Multimodal fusion with bert and attention mechanism for fake news detection. *2021 RIVF Int. Conference on Computing and Communication Technologies (RIVF)*, 1–6.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Vempala, A., & Preo¸tiuc-Pietro, D. (2019). Categorizing and inferring the relationship between the text and image of twitter posts. *Proceedings of the 57th annual meeting of the Association for Computational Linguistics*, 2830–2840.

Wang, W., Tran, D., & Feiszli, M. (2020). What makes training multi-modal classification networks hard? *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 12695–12705.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., . . . Rush, A. (2020). Transformers: State-of-the-art natural language processing. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 38–45. https://doi.org/10.18653/v1/2020.emnlp-demos.6

Wu, P. Y., & Mebane Jr, W. R. (2022). Marmot: A deep learning framework for constructing multimodal representations for vision-and-language tasks. *Computational Communication Research*, *4*(1).

Yang, X., Feng, S., Wang, D., & Zhang, Y. (2020). Image-text multimodal emotion classification via multi-view attentional network. *IEEE Transactions on Multimedia*, *23*, 4014–4026.

Zellers, R., Bisk, Y., Farhadi, A., & Choi, Y. (2019). From recognition to cognition: Visual commonsense reasoning. *Proceedings of the IEEE/CVF conf. on computer vision and pattern recognition*, 6720–6731.

Enhancing Emergency Post Classification through Image Information Amplification via Large Language Models

Pablo Giaccaglia*
Politecnico di Milano, DEIB
pablo.giaccaglia@mail.polimi.it

Carlo A. Bono*
Politecnico di Milano, DEIB
carlo.bono@polimi.it

Barbara Pernici
Politecnico di Milano, DEIB
barbara.pernici@polimi.it

ABSTRACT

Real-time information extracted from social media platforms can be highly valuable during emergencies. For example, reports and direct witnesses can help build situational awareness in the early phases of an emergency, with the potential to save lives. However, suitable techniques for selecting relevant data are needed to gather this information from large-scale social media streams and utilize it effectively. Given the multimedia nature of these streams, selection techniques should simultaneously understand textual and image information, as previous studies highlighted. Leveraging recent advances in language and vision models, we propose and evaluate a method working with a homogeneous, text-only representation for the different modalities of social media posts. Experiments on established and novel datasets, including video data, show that the proposed method achieves state-of-the-art performances while providing a highly general and plug-and-play approach to multimodal data filtering.

Keywords

social media, multimodal data filtering, LLMs

INTRODUCTION

Nowadays, the widespread availability of portable devices and mobile connections has made information production and sharing ubiquitous, making social media a valuable source of information for various purposes, from staying informed to gathering a general overview of a phenomenon. Social media can enhance situational awareness during crisis events by highlighting critical situations, damage to things or people, urgent needs, and other information useful for managing response operations, potentially saving lives. This information often comes in the form of multimedia, such as images and videos. However, the large volume of items posted during crisis events requires proper resources and technologies to rapidly isolate and utilize this information, since their amount is exceptionally high in the first hours following the onset of an emergency.

Several past works investigated the usefulness of social media data for obtaining actionable information during an emergency. The use of social media for collecting up-to-date descriptions during emergencies has been widely advocated over the last decade (Stollberg and De Groeve, 2012), often utilizing a "human as a sensor" paradigm (Avvenuti et al., 2016). Many studies aim to classify social media content according to different categorizations of interest, leveraging various forms of data, including text and multimedia content such as images and videos, to obtain a timely description of ongoing events. Classifying which content is relevant for a given categorization is perceived as a principal barrier to exploiting social media data (Stieglitz et al., 2018).

*corresponding authors

In the present study, we aim to contribute to the task of understanding the informativeness of social media content. Specifically, we look at how augmenting multimodal social media posts with text descriptions derived from Large Language Models (LLMs) might improve their informativeness. We derive textual descriptions from photographic scenes using automatic text generation and combine this additional information with the original text of the posts. Our methodology is tested against two relevant scenarios: a well-known multimodal Twitter dataset (Alam et al., 2018b) and a manually labeled dataset of videos acquired from Reddit. We compare our approach with existing studies and conduct an explainability analysis to understand the nature of the additional information introduced.

The remainder of this work is structured as follows. We summarize literature relevant to our research in the *Related Work* section. Our proposed approach for classifying social media posts is detailed in the *Method* section. In the homonymous sections, we provide the details of the *Datasets* utilized, the *Experiments* performed, and the *Results* obtained by the empirical evaluation of the method. The findings are then examined in the *Discussion* section. Finally, *Conclusions* are drawn in the last section, together with future directions.

RELATED WORK

The prevalent approaches for classifying disaster-related social media posts are grounded in supervised learning. One of the classic works in literature (Sakaki et al., 2013) proposes an algorithm for monitoring tweets to detect earthquake events, focusing on the timeliness of the detection. Over time, approaches based on deep neural networks gained attention (Wiegmann et al., 2020), and production-level platforms for the automatic extraction of relevant posts have been proposed (Imran et al., 2014; Havas et al., 2017). The ability of supervised methods to generalize, especially among different types of events, is one of the critical obstacles, highlighting the need for general-purpose approaches, such as few-shot models (Wiegmann et al., 2020; Kruspe et al., 2020). This challenge also holds for works related to classifying image data or image and text data together (Imran et al., 2020). Recent technological advances have widely upgraded the discerning of both textual and image data. In the context of emergency management, Convolutional Neural Networks (CNNs) have been employed for content classification and damage assessment. The relevance of images during crisis events has been highlighted and exploited by many studies (Peters and De Albuquerque, 2015; Alam et al., 2018a; Bono et al., 2022; Nguyen et al., 2017).

Multimodal fusion techniques aim to enhance classification performance by combining information from both visual and textual modalities (Ramachandram and Taylor, 2017). In recent years, these approaches have been broadly evaluated in emergency scenarios (Abavisani et al., 2020). The integration of modalities can be categorized into two primary strategies. The first and most prevalent involves using distinct backbone networks for each modality, with multimodal fusion occurring either at the classification score level (Wang et al., 2021) or through the combination of high-level features from each network (Kiela et al., 2018; Mozannar et al., 2018), using techniques like addition, outer product, cross-gating, and tensor fusion. The primary limitation of this strategy is the minimal interaction between the different modalities. The second strategy focuses on merging the mid-level features of each modality for a more detailed inter-modal interaction (Huang et al., 2020). Notably, recent multimodal BERT models incorporate transformer layers over these mid-level features, enhancing a fine-grained interplay between text and image features through the attention mechanism (Li et al., 2019). Finally, Liang et al., 2022 propose a method for adapting pre-trained unimodal models like DenseNet and BERT to process image-text pairs for multimodal recognition by integrating cross-modal features without altering the original network structures.

Our study takes a distinct approach from previous ones. Rather than customizing an architecture for combining multimodal data, we convert image information into text with a recent large multimodal model. We then use text-only representations to train a transformer-based classifier. Furthermore, our method is tested with both image and video content, applying it to practical scenarios and including a type of media not frequently considered in similar studies.

METHOD

In the present work, we investigate whether automatically derived textual content, such as captions and descriptions obtained from LLMs, can support multimedia classification and improve the ability to distinguish informative content published on social media. We utilize Large Language and Vision Assistant (LLaVA) (H. Liu et al., 2023), a state-of-the-art large multimodal model, for generating image-related descriptions, and Robustly Optimized BERT Approach (RoBERTa) (Y. Liu et al., 2019) for training the classifier on the resulting representation. We focus on a binary classification setup, aiming at differentiating "*informative*" and "*not informative*" content, according to a context-dependent definition of informativeness. The generated text is concatenated to the original text of the post, thus obtaining a homogeneous, augmented textual representation. A RoBERTa classifier is then trained over this representation. The architecture of the proposed method is illustrated in Figure 1.

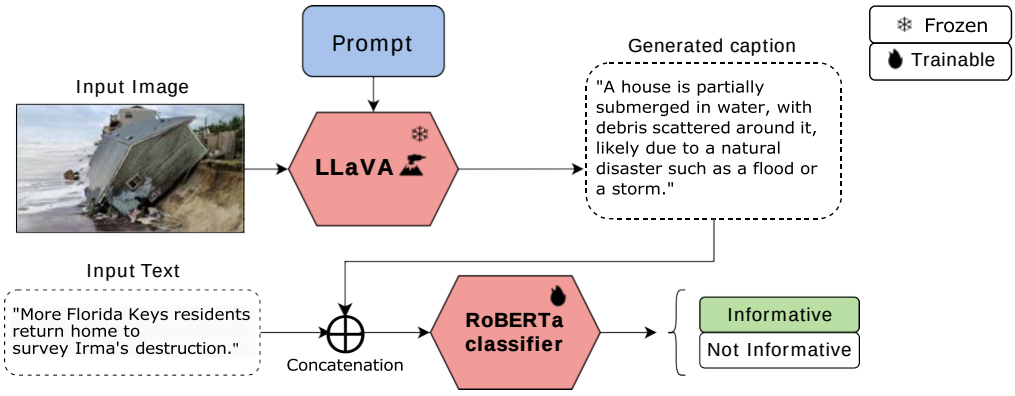


Figure 1. **Architecture of the proposed integration and classification method**

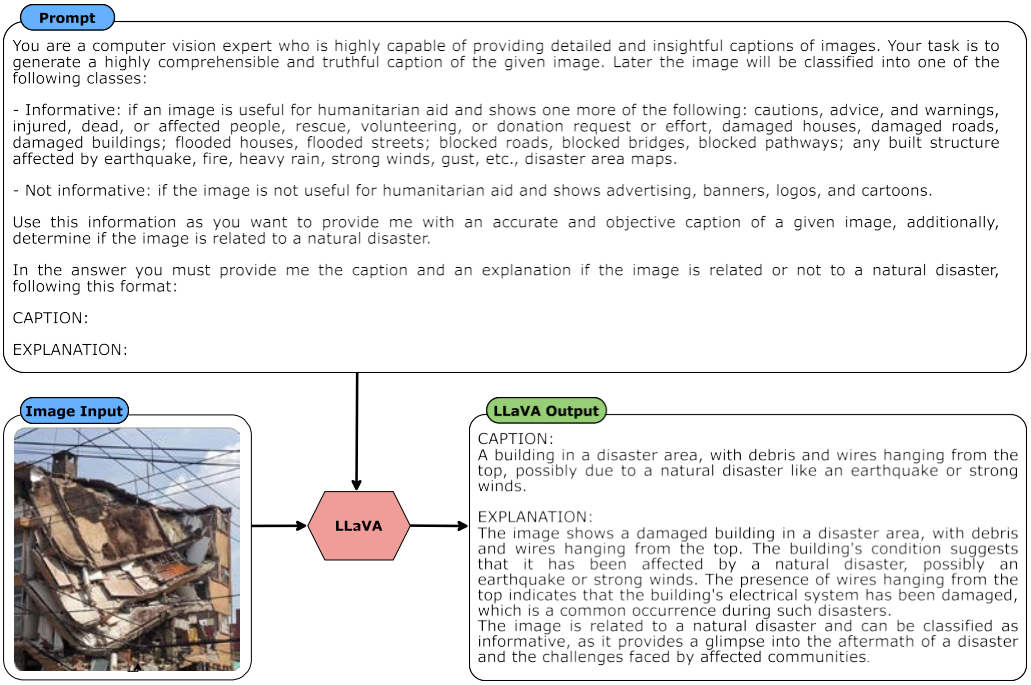


Figure 2. **Information augmentation for the**

As discussed in the previous section, fusion models usually require ormation from text and images, as these data modalities have inherently different hallenges in model design and optimization. When employing image-only models, visual features, mainly disconnected from the semantic features of textual data. use of a text-only classifier, avoiding the need for more complex fusion methods t performance.

LLaVA is a large, general-purpose multimodal model combining visual and a vision encoder and the Vicuna chat assistant (Zheng et al., 2023). The model information from images according to a preferred semantic level, such as an imag or task-dependent reasoning. A fundamental feature of the approach is its ability textual ones, allowing for control over the semantic level and format of the generated prompt engineering.

In the considered setup, multiple text generation tasks are requested, such as visual-onl The idea is to exploit the language model, through suitable prompts, for different kinds of which could in turn improve the downstream classification task. An example is reported in 2 detailed in the Experiments section. Each component of the generated information, derived frames attached to posts, can be regarded as supplementary text, extending the original pos

Table 1. **Characteristics of the datasets**

	Number of posts	Number of images	*Informative* images (%)
CrisisMMD	11,400	12,708	66.63
Reddit	838	35,551	6.39

Each type of resulting textual representation is then separately fed to a transformer-based model for training and evaluation.

DATASETS

We conducted tests on two distinct datasets to benchmark the effectiveness of the proposed method. The first one, CrisisMMD, is a well-established multimodal dataset collected from Twitter during natural emergencies (Alam et al., 2018b). The second is a novel dataset consisting of video content collected from Reddit over one week at the onset of the 2023 conflict between Israel and Hamas.

CrisisMMD The dataset was collected using event-specific keywords and hashtags during seven natural disasters that occurred in 2017: Hurricane Irma, Hurricane Harvey, Hurricane Maria, the Mexico earthquake, the California wildfires, the Iraq-Iran earthquakes, and the Sri Lanka floods. The dataset has been annotated for three tasks. In the current study, we specifically focus on the *informativeness* task, which aims at recognizing whether a social media post is informative or uninformative for humanitarian aid purposes (see the definition of *informativeness* in Ofli et al., 2020).

In CrisisMMD, text and image pairs, which are annotated separately, could carry different labels. To properly compare the proposed method with previous works (Liang et al., 2022; Ofli et al., 2020), we align with the benchmark guidelines by focusing exclusively on the portion of the dataset where text and images share the same label for the informativeness task. The resulting dataset composition is reported in Table 1.

Reddit Videos To evaluate our method on a recent event, we collected Reddit posts with video attachments during the initial days of the 2023 conflict between Israel and Hamas. We used the Reddit API via the *praw*[1] library, using "*gaza OR israel OR hamas OR palestin**" as a case-insensitive search query, gathering 49,788 posts from October 13th to 19th, 2023. To account for the different nature of this dataset, we adapted the definition of *informativeness* as follows. The post is considered *informative* if it reports or shows direct proof of damage to people, such as injured, dead, or affected persons, or damage to man-made things, such as buildings, infrastructures, or vehicles, including the damaging act itself. Non-primary sources and indirect reporting, including newscasts, interviews performed either in a studio or in the field, reaction videos, edited videos, and non-photographic videos, are considered *not informative*.

Among the collected posts, 2,644 had a video attachment. Of these posts, 1,959 had a text length of at least ten words. This minimum word count was chosen arbitrarily to focus on posts with substantial descriptions. We filter out duplicated videos via video hashing[2].

Videos are then classified as *informative* based on the definition mentioned above, with two annotators using Label Studio[3]. We only keep videos where both annotators agreed on the label. Analogously to the data selection performed in the previous dataset, the annotators performed a second pass to exclude a limited number of videos for which the labels and the original posts' text did not agree. Then, we extracted all the keyframes from the videos and eliminated duplicated frames using image hashing[4]. Each frame was assigned the label and text from its original post, although we recognize that this method might not be perfectly accurate. The final dataset composition is reported in Table 1. This dataset, which includes some level of noise and is related to a human-made emergency, is meant to test the method's effectiveness under challenging conditions[5].

1https://github.com/praw-dev/praw
2https://pypi.org/project/videohash/
3https://labelstud.io/
4https://github.com/jgraving/imagehash
5The dataset with submission IDs and annotations is available upon request

EXPERIMENTS

LLM Setup

Regarding the text augmentation, we used LLaVA 13B version 1.5 for the CrisisMMD task and LLaVA 7B version 1.5 for the Reddit task because of budget limitations. These models were trained in September 2023 on various datasets, including a set of GPT-generated multimodal instruction-following data. The architecture adopted for the text classifier is RoBERTa base6, pre-trained on 58M tweets and fine-tuned for sentiment analysis with the TweetEval benchmark.

For the text augmentation on the CrisisMMD dataset, the LLaVA *Temperature* parameter, which controls the randomness of the output, was set to 0.7, while the *Maximum Number of Generated Tokens* was set to 1024. The *Nucleus Sampling* parameter, also known as *Top P*, which determines diversity by only considering a subset of the most probable next words, was set to 0.7. In the case of text augmentation on the Reddit dataset, the settings used were analogous to those for the CrisisMMD dataset, except for the *Temperature* parameter, which was adjusted to 0.2. This is done due to the nature of the dataset, which frequently features video frames with significant motion and complexity. By lowering the *Temperature* setting, the model's range of generative options is somewhat limited, which supposedly helps to generate consistent descriptions of the video frames.

Classifier Setup

LLaVa weights are kept frozen and used only for text generation. The weights of the text classifier are instead fine-tuned while training for the informativeness classification task. We use Adam as optimizer, using a mini-batch size of 64, and weighted cross-entropy as a loss function, with weights accounting for class imbalance. We employ a Cyclical Learning Rate policy (Smith, 2017) for mini-batch Gradient Descent. Cyclical Learning Rate adjusts the learning rate by cyclically changing it between a minimum value of $1e^{-5}$ and a maximum value of $1e^{-4}$. To mitigate overfitting, we use an early-stopping condition on the validation cross-entropy loss, with patience of 5 epochs. Both LLaVa generation and RoBERTa training were performed on an A40 GPU. We did not specifically tune any hyper-parameter; hence, there is likely room for improvements in future studies.

Models are assessed using three metrics: classification accuracy, macro-averaged F1-score (M-F1), and weighted F1-score (W-F1), with weights determined by the number of samples for each class. It is important to note that during emergency events the distribution of samples across categories is usually imbalanced. This motivated the choice of macro-averaged and weighted F1-score to properly synthesize false positives and false negatives.

CrisisMMD

To compare our method against previous works (Liang et al., 2022), we keep the same training, validation, and test data splits provided by the dataset creators (Alam et al., 2018b). Table 2 shows the total number of tweets and attached images for each category, along with their distribution among training, validation, and test sets. The variation in the number of tweets and images in the training set is due to some tweets having multiple image attachments. These tweets are assigned exclusively to the training set.

Table 2. **Labels and data splits for the CrisisMMD informativeness task**

	Train (70%)		**Valid (15%)**		**Test (15%)**		**Total**	
	Text	**Image**	**Text**	**Image**	**Text**	**Image**	**Text**	**Image**
Informative	5,546	6,345	1,056	1,056	1,030	1,030	7,632	8,431
Not informative	2,747	3,256	517	517	504	504	3,768	4,277
Total	8,293	9,601	1,573	1,573	1,534	1,534	11,400	12,708

We used LLaVA to generate textual descriptions for images using prompt engineering techniques. We instructed the model to act as a computer vision expert and outlined the specific tasks. These tasks included generating a descriptive caption and an explanation for a given image, with information related to the specific classification task. Additionally, we specified the format for the two required outputs: a "*CAPTION*" providing a concise and factual image description and an "*EXPLANATION*" justifying if and why a given image could be *informative* in the current scenario. The prompt utilized is reported in Figure 2.

6https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment

Table 3. **Labels and data splits for the Reddit informativeness task**

	Train & Valid (70% - 15%)		Test (15%)		Total	
	Text	**Frame**	**Text**	**Frame**	**Text**	**Frame**
Informative	61	2,066	11	208	72	2,274
Not informative	651	28,490	115	4,787	766	33,277
Total	712	30,556	126	4,995	838	35,551

We processed the generated texts to craft six distinct inputs for the classifier, namely: the tweet text concatenated with the caption; the tweet text concatenated with the explanation; The tweet text concatenated with both the caption and the explanation; the caption concatenated with the explanation, without the tweet text; the caption alone; the explanation alone.

We compared our approach with traditional networks trained on a single modality (unimodal image and text, Ofli et al., 2020), as well as existing state-of-the-art image-text multimodal classification methods. Studies such as Abavisani et al., 2020 primarily employ a multimodal fusion of global features from each unimodal backbone. Others, like Huang et al., 2020, employ pre-trained multimodal BERT models adapted for image-text classification and inter-modal interaction via the attention mechanisms. We also compare to standard multimodal recognition techniques: score fusion (averaging unimodal classification scores) and feature concatenation (merging global features from each unimodal source). The study from Liang et al., 2022, which proposes a method for integrating information from other modalities into the unimodal models, was considered as a reference for the reported performances of the cited models.

Reddit Videos

Similarly to the CrisisMMD task, we used LLaVA to obtain information about the video frames. We instructed the model to act as a computer vision expert analyzing video frames. The requested output consisted of three text elements: a "*CAPTION*" providing a brief, factual description of the image; an "*EXPLANATION*" offering a justification aligned with the informativeness task; and a "*REASONING*" section, which elucidates the model's "reasoning" concerning the content of the image and its connection to the classification task described. The prompt utilized is reported in Figure 3.

We processed the generated texts to craft six distinct inputs for the classifier, namely: the post text concatenated with the caption; the post text concatenated with the explanation; the post text concatenated with the reasoning; the post text concatenated with both the caption and the explanation; the caption alone; the explanation alone. Finally, the dataset was partitioned into training, validation, and test sets, accounting for 70%, 15%, and 15% of the posts, respectively. We ensured, by leveraging video identifiers, that all frames from the same video appeared only in one partition. When partitioning, we ensured the preservation of class balance among the partitions by stratifying based on class labels. Table 3 displays the total posts and video frames for each category, including their distribution over splits.

To evaluate the robustness and generalization ability of the proposed method on the Reddit dataset, we replicated the same experiment over five random samples. We conducted the same experiment across five distinct pairs of training and validation sets, each generated through random partitioning of non-test data. We ensured that frames from the same video appeared only in one fold, and stratified the sampling on class labels. To evaluate the classifier's performance also at the video level, a video is classified as *informative* if at least 50% of its frames are predicted as *informative*.

RESULTS

CrisisMMD

In Table 4, we report the performance results achieved by different methods, including the proposed one, on the informativeness task.

In Figure 4, we present the test set confusion matrices for the multimodal model developed by Ofli et al., 2020 and for the three best-performing models proposed in this work. Confusion matrices from other studies are not included, as they were not provided in the respective papers.

[†]Reported performances are taken from Liang et al., 2022

[*]Computed from the corresponding confusion matrix reported by Ofli et al., 2020

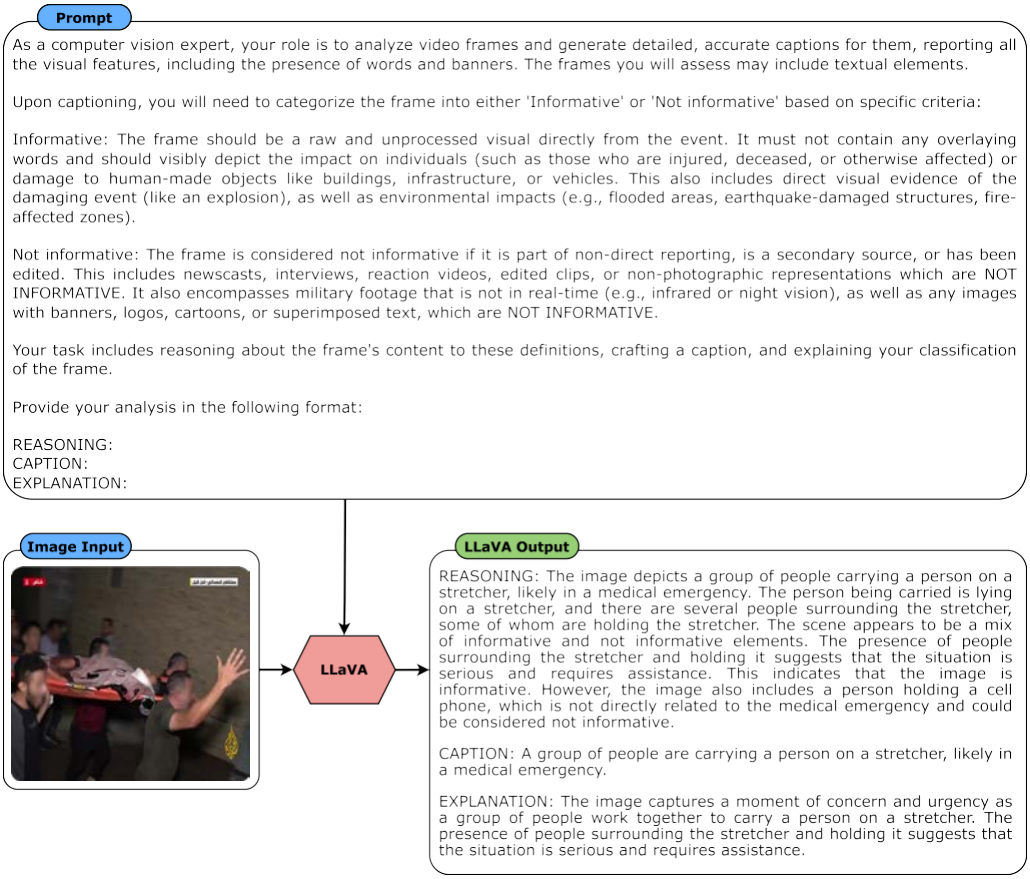


Figure 3. **Information augmentation for the Reddit informativeness task**

Table 4. **Results for CrisisMMD informativeness task**

Method	Acc	M-F1	W-F1
Unimodal Text (Ofli et al., 2020)	80.8	78.5*	80.9
Unimodal Image (Ofli et al., 2020)	83.3	80.5*	83.2
Multimodal Fusion (Ofli et al., 2020)	84.4	81.9*	84.2
Cross-attention (Abavisani et al., 2020)†	88.4	87.6	88.7
CentralNet (Vielzeuf et al., 2018)†	87.8	85.3	86.1
GMU (Ovalle et al., 2017)†	87.2	84.6	85.7
CBP (Fukui et al., 2016)†	87.9	85.6	86.4
CBGP (Kiela et al., 2018)†	88.1	86.7	87.3
MMBT (Kiela et al., 2019)†	86.4	85.3	86.2
VisualBERT (Li et al., 2019)†	88.1	86.7	88.6
PixelBERT (Huang et al., 2020)†	88.7	86.4	87.1
ViLT (Kim et al., 2021)†	87.6	85.1	88.0
ME Feature Concat (Liang et al., 2022)†	90.8	**91.6**	90.3
ME Cross-attention (Liang et al., 2022)†	**92.0**	91.2	91.3
RoBERTa Text Only	86.6	84.8	86.6
LLaVa+RoBERTa Caption Only	85.7	82.8	85.2
LLaVa+RoBERTa Explanation Only	85.5	83.3	85.4
LLaVa+RoBERTa Caption and Explanation	86.8	84.8	86.7
LLaVa+RoBERTa Text & Caption	91.4	90.2	**91.4**
LLaVa+RoBERTa Text & Explanation	88.8	87.6	88.9
LLaVa+RoBERTa Text & Caption & Explanation	90.6	89.3	90.6

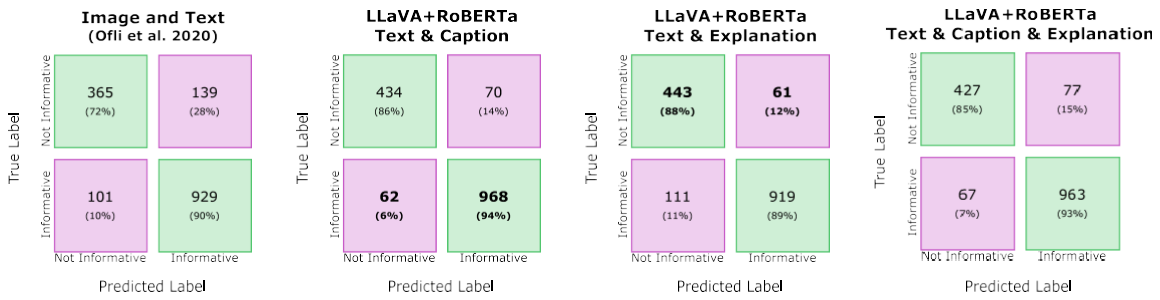


Figure 4. **Test set confusion matrices for the CrisisMMD informativeness task**

Table 5. **Results for Reddit informativeness task**

Method	Acc	W-F1	*Informative* F1
Frame metrics (5 folds average and std)			
LLaVA+RoBERTa Only Text	86.2 (1.0)	88.0 (1.0)	40.0 (6.8)
LLaVA+RoBERTa Text & Caption	96.3 (0.7)	96.3 (0.6)	54.1 (6.7)
LLaVA+RoBERTa Text & Reasoning	96.0 (0.4)	96.0 (0.6)	47.2 (10)
LLaVA+RoBERTa Text & Explanation	96.1 (1.1)	96.0 (0.9)	52.0 (8.8)
LLaVA+RoBERTa Text & Caption & Explanation	95.3 (2.8)	95.6 (2.0)	52.0 (9.0)
LLaVA+RoBERTa Only Caption	89.0 (0.9)	91.4 (0.5)	34.5 (2.0)
LLaVA+RoBERTa Only Explanation	89.4 (1.8)	91.7 (1.0)	34.0 (1.9)
Video metrics - 0.5 threshold (5 folds average and std)			
LLaVA+RoBERTa Text & Caption	90.1 (1.3)	90.0 (1.3)	40.0 (8.0)
LLaVA+RoBERTa Text & Reasoning	90.1 (0.6)	90.0 (0.7)	40.0 (6.0)
LLaVA+RoBERTa Text & Explanation	**90.5** (1.3)	90.1 (0.9)	41.3 (3.0)
LLaVA+RoBERTa Text & Caption & Explanation	89.2 (3.4)	89.2 (2.0)	39.6 (7.8)
LLaVA+RoBERTa Only Caption	90.0 (2.0)	**91.0** (1.3)	**60.0** (3.0)
LLaVA+RoBERTa Only Explanation	90.0 (2.15)	91.0 (1.6)	53.4 (5.6)

Reddit Videos

In Table 5, we report the performance results achieved by the proposed method on the Reddit informativeness task. The table reports the performance metrics computed both at the frame and video levels.

DISCUSSION

CrisisMMD

We can draw the following observations from Table 4. First, the unimodal models generally perform worse than the multimodal approaches. Second, we can see that the late unimodal fusion models (Cross-attention, CentralNet, GMU, CBP, CBGP), as well as the large pre-trained multimodal BERT models (MMBT, VisualBERT, PixelBERT, and ViLT) share similar performances.

The proposed approach, when injecting LLM-generated information, achieves state-of-the-art performances. We observe that the models that consider the original texts of the tweets generally perform better than the ones using only LLM-generated information. Although expected, this highlights the relevance of the tweets' contents for discerning relevance. In addition, we observe that the model combining the original text and captions performs better than combining the original text with other LLM-generated information. This fact may be due to the brief, factual information provided by the captions, which are possibly more helpful to the classifier. In contrast, explanations tend to be more detailed, including descriptions, comments, and evaluations. These additional details could lead the classifier astray, making it possibly more complicated to interpret the overall text accurately.

The confusion matrices presented in Figure 4 show that the proposed model combining tweet text and captions significantly outperforms the other three models in correctly identifying true positives and minimizing false negatives. Although this configuration generates a few more false positives than the model that combines text and explanations, a higher number of false negatives would be far more problematic during an emergency than a slight increase in false positives. The performance looks slightly worse when text is paired with explanations than when it

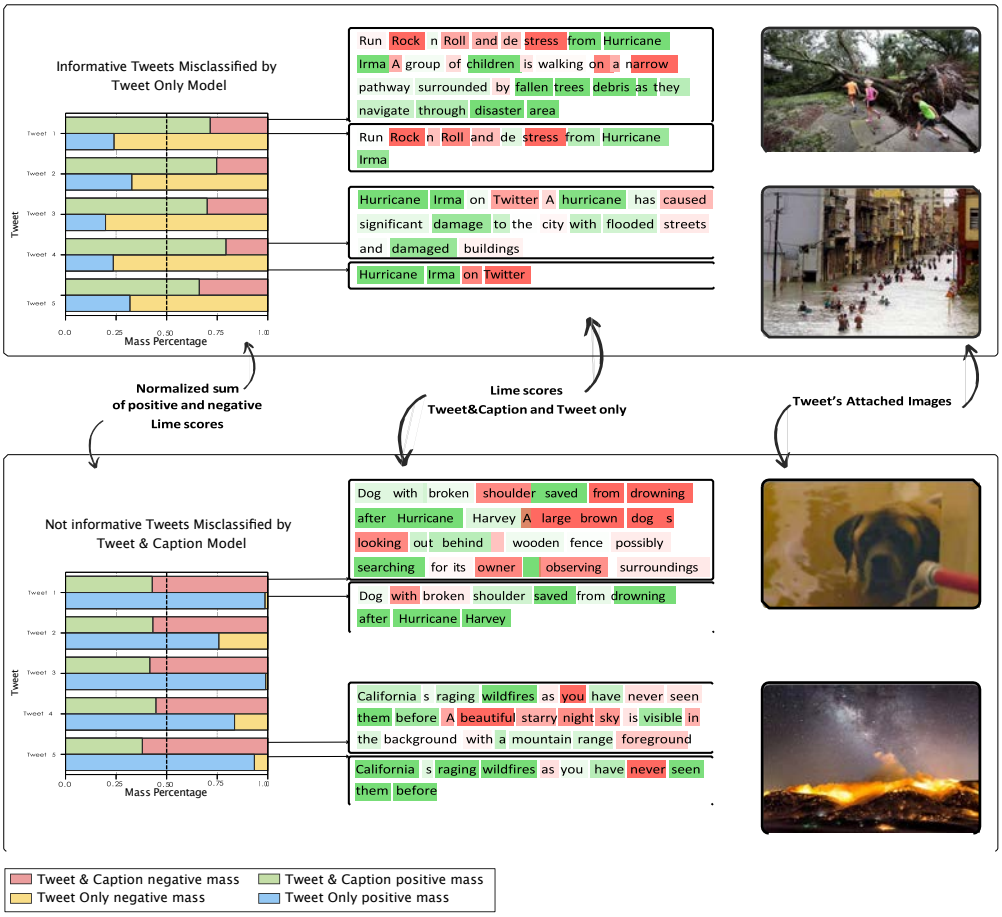


Figure 5. **Information Amplification Analysis in CrisisMMD: On the left, normalized LIME scores are displayed for selected tweets; in the middle, word-level LIME scores are highlighted (green for positive, red for negative); on the right, the images linked to each tweet are reported. *Positive* and *negative* mass percentages refer to the normalized sum of LIME scores within each tweet**

is paired with captions. When text, captions, and explanations are combined, the performance falls between the two preceding scenarios, possibly due to the beneficial effect of captions. This supports the hypothesis that explanations make the task more difficult in this scenario.

To understand the effect of the information amplification performed by the LLM-generated text, we performed an explainability analysis via Local Interpretable Model-Agnostic Explanation (LIME, Ribeiro et al., 2016). LIME perturbates the input instances and uses the model's predictions on these instances to fit an interpretable model in the proximity of the selected instance; this second model reveals which features are influential for the specific instance and how much they contribute to the global model's prediction.

Figure 5 illustrates how the LLM augmentation can help the classification task. On the left, each pair of bars represents an *informative* Twitter post from the CrisisMMD dataset, with the different areas representing the normalized sums of positive and negative LIME scores for two configurations, *"Tweet & Caption"* (first row) and *"Tweet only"* (second row).

In the upper section, we observe that positive scores are boosted after augmenting the text with LLM-generated captions, recovering false negative cases. However, the information injection also has a counterproductive effect in some cases, as seen in the bottom section. Here, examples of *informative* tweets are misclassified by the *Tweet & Caption* model due to features with negative scores becoming prevalent. Although the balance between misclassifications remains favorable compared to not using information amplification, as shown in Table 4, this observation highlights the challenges in managing the impact of the additional information introduced.

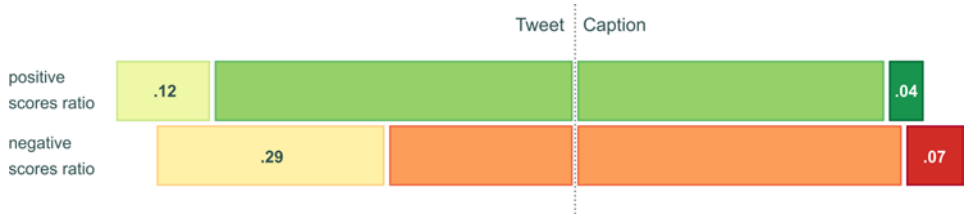


Figure 6. **Normalized LIME score ratios in CrisisMMD test set; Tweet-only lemmata on the left; Caption-only lemmata on the right; lemmata shared by both sections in the middle**

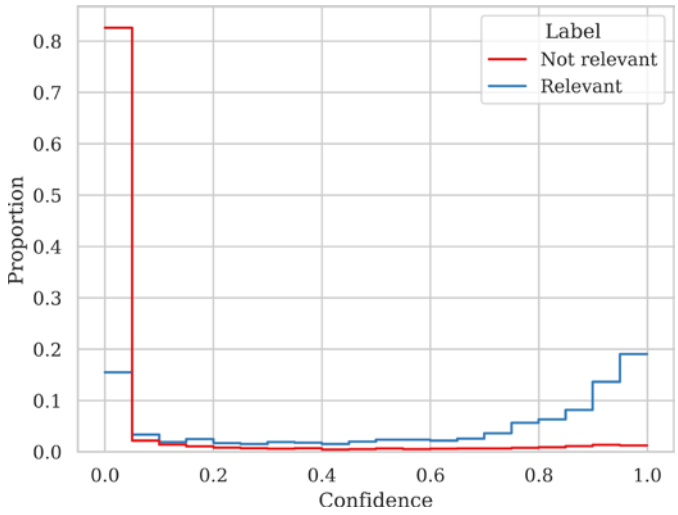


Figure 7. **Test set frame confidence distribution, normalized per class (Reddit Task)**

Figure 6 provides a visual representation of the overall effect of the information introduced by the LLM. The figure refers to the *Tweet & Caption* model and highlights the contribution of each of the two sources – original tweet and caption – to the classification outcome by aggregating and normalizing the LIME scores related to lemmata[7], after excluding English stopwords. In the figure, positive and negative scores are aggregated as rows, while a vertical line partitions tweet-dependent and caption-dependent scores. The three colors highlight the characteristics of the lemmata: the two lighter blocks on the left account for the weight of lemmata seen only in tweets, while the two darker blocks on the right account for the weight of lemmata seen only in captions.

As a general observation, it can be seen that most of the scores originate from lemmata shared between the two sources. The effect of new terms introduced by the LLM is limited, roughly one-fourth of the percentage of scores for tweet-only lemmata. This observation suggests that the tweet's text could consistently describe the images' contents. Overall, the original tweet text accounts for 57% of the positive scores and 52% of the negative scores.

Reddit Videos

The proposed approach shows a good separability of the classes already at the frame level, as the distributions in Figure 7 show. Roughly 16% of the *informative* frames are classified as *not informative* with more than 95% confidence. Upon manual inspection of a sample of these frames, the reasons behind the misclassifications were not completely clear, since some relevant frames were present in this subset, together with generic, blank, and indistinct frames isolated from videos labeled as *"informative"*. This observation also suggests that a more fine-grained approach to labeling, such as considering video segments instead of whole videos, could be beneficial for the task, albeit more burdensome as an annotation task. On the other hand, this does not impair the performance of the video classifier, since only individual frames are affected by the misclassification.

Moreover, it should be noted that the reported performance results, based on common-sense thresholds — 0.5 confidence at the frame level and 0.5 proportion of positive frames at the video level — prove to be conservative

[7] "*A form of a word that appears as an entry in a dictionary and is used to represent all the other possible forms*" according to the Cambridge Dictionary

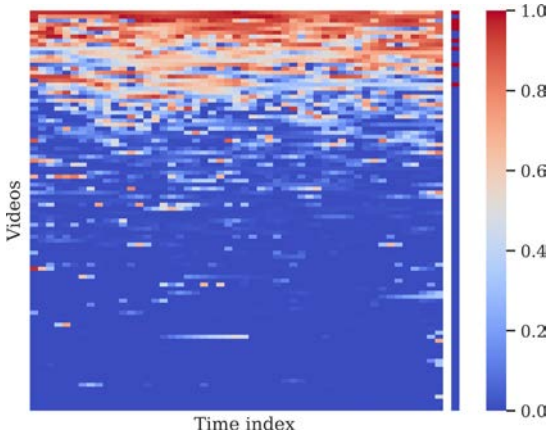


Figure 8. **Confidence for "*informative*" class for each test set video, over normalized time index. True label on the right (Reddit Task)**

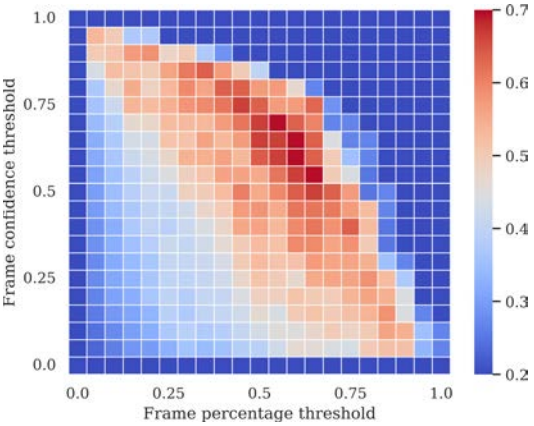


Figure 9. **Response of the test set F1-score for the "*informative*" class varying classification thresholds (Reddit Task)**

estimates of the performance. Figure 9 reports a grid exploration of the thresholds, showing that an increase in the video-level threshold could lead to better performances. We have chosen to evaluate our approach using a common-sense threshold to assess its performance as a general-purpose method. It should also be noted that the *"Caption Only"* model performs better at the video level, even if the F1 score at the frame level is worse. This is due to a higher false positive rate in the majority class (*"Not Informative"*), leading to a better final result since frames predicted as positive are still a minority when considered at the video level.

An understanding of the frame-level labels for the test set videos can be obtained from Figure 8. The figure illustrates how the confidence of being an *informative* frame is distributed along the videos, with red indicating higher confidence. The video durations are normalized for convenience. It can be observed that higher confidence regions, while still present in many of the *not informative* videos (blue region in the bottom part of the graph), are scattered across the video length; these false positives are mostly absorbed by the threshold at the video level. The fact that the *"Caption Only"* performs reasonably well without relying on the original text could also be related to the fact that, often, the text of the Reddit submission in the scenario analyzed is not merely descriptive but contains additional considerations, such as stances and advocacy.

CONCLUSIONS AND FUTURE WORK

We proposed a method for categorizing multimodal posts by extracting text-only representations from multimedia attachments and merging this information with the text of social media posts. One of the advantages of this approach is that it is general-purpose and we showed that it achieves state-of-the-art performances on the informativeness task without the need of fine tuning for a specific case. Further room for improvement is left by performing hyper-parameter tuning on the classifier on the specific case, although this was not the goal of the present study.

Amalgamating text and image information in social media data is crucial, as they often present complementary or conflicting information. Since current multimodal classification methods assume common labels for different modalities, future works should investigate multimodal learning algorithms capable of handling heterogeneous input (e.g., text and image pairs with conflicting labels).

One of the main hindrances to adopting LLMs is their computational cost, which is linked to energy utilization and environmental consequences. We plan to examine the trade-offs offered by alternative smaller models and efficiency-oriented methods such as network quantization, considering the combined effect of performance, execution times, and computational requirements to support a wider choice of use cases.

Finally, we plan to extend the evaluation of the proposed approach to few-shot and continual learning scenarios, since the intrinsically dynamic nature of emergency events poses significant challenges to adopting relevance-oriented classifiers. Moreover, since historical emergency-related data can be limited depending on the kind of emergency, learning informativeness on the fly could be the only viable option in some scenarios.

ACKNOWLEDGEMENTS

This work has been supported by the PNRR-PE-AI "FAIR" project funded by the NextGenerationEU program and by the PRIN 2022 Project "Discount quality for responsible data science: Human-in-the-Loop for quality data".

REFERENCES

Abavisani, M., Wu, L., Hu, S., Tetreault, J., & Jaimes, A. (2020). Multimodal categorization of crisis events in social media. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14679–14689.

Alam, F., Ofli, F., & Imran, M. (2018a). Processing social media images by combining human and machine computing during crises. *International Journal of Human–Computer Interaction*, *34*(4), 311–327.

Alam, F., Ofli, F., & Imran, M. (2018b). CrisisMMD: Multimodal Twitter datasets from natural disasters. *Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM)*.

Avvenuti, M., Cimino, M. G. C. A., Cresci, S., Marchetti, A., & Tesconi, M. (2016). A framework for detecting unfolding emergencies using humans as sensors. *SpringerPlus*, *5:43*.

Bono, C., Pernici, B., Fernandez-Marquez, J. L., Shankar, A. R., Mu¨laˆyim, M. O., & Nemni, E. (2022). TriggerCit: Early Flood Alerting using Twitter and Geolocation–a comparison with alternative sources. *Proc. ISCRAM 2022, Tarbes, France*.

Fukui, A., Park, D. H., Yang, D., Rohrbach, A., Darrell, T., & Rohrbach, M. (2016). Multimodal compact bilinear pooling for visual question answering and visual grounding. In J. Su, X. Carreras, & K. Duh (Eds.), *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP), Austin, Texas, USA* (pp. 457–468). The Association for Computational Linguistics.

Havas, C., Resch, B., Francalanci, C., Pernici, B., Scalia, G., Fernandez-Marquez, J. L., Van Achte, T., Zeug, G., Mondardini, R., Grandoni, D., Kirsch, B., Kalas, M., Lorini, V., & Ru¨ping, S. (2017). E2mC: Improving emergency management service practice through social media and crowdsourcing analysis in near real time. *Sensors*, *17*(12). https://doi.org/10.3390/s17122766

Huang, Z., Zeng, Z., Liu, B., Fu, D., & Fu, J. (2020). Pixel-BERT: Aligning image pixels with text by deep multi-modal transformers. *arXiv preprint arXiv:2004.00849*.

Imran, M., Castillo, C., Lucas, J., Meier, P., & Vieweg, S. (2014). AIDR: Artificial. *Proceedings of the 23rd International Conference on World Wide Web*, 159–162. https://doi.org/10.1145/2567948.2577034

Imran, M., Ofli, F., Caragea, D., & Torralba, A. (2020). Using AI and social media multimodal content for disaster response and management: Opportunities, challenges, and future directions. *Information Processing & Management*, *57*(5), 102261.

Kiela, D., Bhooshan, S., Firooz, H., & Testuggine, D. (2019). Supervised multimodal bitransformers for classifying images and text. *Visually Grounded Interaction and Language (ViGIL), NeurIPS 2019 Workshop, Vancouver, Canada, December 13, 2019*.

Kiela, D., Grave, E., Joulin, A., & Mikolov, T. (2018). Efficient large-scale multi-modal classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, *32*(1), 5198–5204.

Kim, W., Son, B., & Kim, I. (2021). ViLT: Vision-and-Language Transformer without convolution or region supervision. *International Conference on Machine Learning*, 5583–5594.

Kruspe, A., Kersten, J., & Klan, F. (2020). Detecting event-related tweets by example using few-shot models. *ISCRAM 2019 Conference Proceedings – 16th International Conference on Information Systems for Crisis Response and Management*, 825–835.

Li, L. H., Yatskar, M., Yin, D., Hsieh, C.-J., & Chang, K.-W. (2019). VisualBERT: A simple and performant baseline for vision and language. *arXiv preprint arXiv:1908.03557*.

Liang, T., Lin, G., Wan, M., Li, T., Ma, G., & Lv, F. (2022). Expanding large pre-trained unimodal models with multimodal information injection for image-text multimodal classification. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 15492–15501.

Liu, H., Li, C., Wu, Q., & Lee, Y. J. (2023). Visual instruction tuning. *NeurIPS*.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692*.

Mozannar, H., Rizk, Y., & Awad, M. (2018). Damage identification in social media posts using multimodal deep learning. *International Conference on Information Systems for Crisis Response and Management*.

Nguyen, D. T., Ofli, F., Imran, M., & Mitra, P. (2017). Damage assessment from social media imagery data during disasters. *2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 569–576.

Ofli, F., Alam, F., & Imran, M. (2020, May). Analysis of social media data using multimodal deep learning for disaster response. In A. L. Hughes, F. McNeill, & C. W. Zobel (Eds.), *17th International Conference on Information Systems for Crisis Response and Management (ISCRAM)* (pp. 802–811). ISCRAM Digital Library.

Ovalle, J. E. A., Solorio, T., Montes-y-Go´mez, M., & Gonza´lez, F. A. (2017). Gated multimodal units for information fusion. *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Workshop Track Proceedings*.

Peters, R., & De Albuquerque, J. P. (2015). Investigating images as indicators for relevant social media messages in disaster management. *12th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*.

Ramachandram, D., & Taylor, G. W. (2017). Deep multimodal learning: A survey on recent advances and trends. *IEEE Signal Processing Magazine*, *34*(6), 96–108. https://doi.org/10.1109/MSP.2017.2738401

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Sakaki, T., Okazaki, M., & Matsuo, Y. (2013). Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Trans. Knowl. Data Eng.*, *25*(4), 919–931. https://doi.org/10.1109/TKDE.2012.29

Smith, L. N. (2017). Cyclical learning rates for training neural networks. *2017 IEEE Winter Conf. on Applications of Computer Vision (WACV)*, 464–472.

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics – challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*, *39*, 156–168. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2017.12.002

Stollberg, B., & De Groeve, T. (2012). The use of social media within the global disaster alert and coordination system (GDACS). *Proc. 21st Intl. WWW Conf.*, 703–706.

Vielzeuf, V., Lechervy, A., Pateux, S., & Jurie, F. (2018). CentralNet: A multilayer approach for multimodal fusion. *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 575–589.

Wang, Y., Xu, X., Yu, W., Xu, R., Cao, Z., & Shen, H. T. (2021). Combine early and late fusion together: A hybrid fusion framework for image-text matching. *2021 IEEE International Conference on Multimedia and Expo (ICME)*, 1–6. https://doi.org/10.1109/ICME51207.2021.9428201

Wiegmann, M., Kersten, J., Klan, F., Potthast, M., & Stein, B. (2020). Analysis of detection models for disaster-related tweets. *ISCRAM 2020 Conference Proceedings – 17th International Conference on Information Systems for Crisis Response and Management*, 872–880.

Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E. P., Zhang, H., Gonzalez, J. E., & Stoica, I. (2023). Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena. *arXiv preprint arXiv:2306.05685*.

Monitoring Critical Infrastructure Facilities During Disasters Using Large Language Models

Abdul Wahab Ziaullah
Qatar Computing Research Institute,
Hamad Bin Khalifa University, Doha, Qatar
awahab@hbku.edu.qa

Ferda Ofli
Qatar Computing Research Institute,
Hamad Bin Khalifa University, Doha, Qatar
fofli@hbku.edu.qa

Muhammad Imran
Qatar Computing Research Institute,
Hamad Bin Khalifa University, Doha, Qatar
mimran@hbku.edu.qa

ABSTRACT

Critical Infrastructure Facilities (CIFs), such as healthcare and transportation facilities, are vital for the functioning of a community, especially during large-scale emergencies. In this paper, we explore a potential application of Large Language Models (LLMs) to monitor the status of CIFs affected by natural disasters through information disseminated in social media networks. To this end, we analyze social media data from two disaster events in two different countries to identify reported impacts to CIFs as well as their impact severity and operational status. We employ state-of-the-art open-source LLMs to perform computational tasks including retrieval, classification, and inference, all in a zero-shot setting. Through extensive experimentation, we report the results of these tasks using standard evaluation metrics and reveal insights into the strengths and weaknesses of LLMs. We note that although LLMs perform well in classification tasks, they encounter challenges with inference tasks, especially when the context/prompt is complex and lengthy. Additionally, we outline various potential directions for future exploration that can be beneficial during the initial adoption phase of LLMs for disaster response tasks.

Keywords

Large language models, disaster management, social media, information classification, information retrieval

INTRODUCTION

Critical infrastructures, including essential facilities and services vital for societal functioning, span sectors like energy, transport, healthcare, and telecommunications (Labaka et al., 2016; Pescaroli & Kelman, 2017). Ensuring the continuous functioning and accessibility of critical infrastructures is essential in assisting vulnerable populations during disasters and disruptions to these critical infrastructures can result in increased human and economic losses (Auerswald et al., 2005). However, during major disasters, authorities are overwhelmed and thus face challenges in maintaining an updated status of these facilities, and as a consequence, the general public is kept uninformed about the latest status of these facilities.

This paper bridges the gap by exploring an unconventional data source—social media—to monitor Critical Infrastructure Facilities (CIFs) in a given area of interest (AOI). Social media platforms, especially microblogging sites such as X (formerly Twitter), are important data sources for real-time disaster updates. People post information about early warnings, cautions, damages, and their needs almost instantaneously (Olteanu et al., 2015; Qu et al., 2011). Past studies show the use of social media information for several disaster response tasks including damage and needs assessment, among others (Castillo, 2016; Imran et al., 2015).

While social media holds valuable information, it frequently becomes overloaded with noise—comprising casual chatter filled with slangs, abbreviations, and grammatically unconventional sentences. Past studies suggest employing supervised classification techniques to discern social media messages relevant to disaster response (Imran et al., 2015). However, these techniques require human-labeled data for each unique classification task or even when adapting a model trained on past disaster data to a new emergency. For example, Alam et al. (2021) manually labeled around 77K tweets to train deep neural network models such as BERT, RoBERTa, for detecting situational awareness messages. Similarly, Zahra et al. (2020) shows an extensive feature engineering process to train classifiers for eyewitness message detection. These methods are tailored for specific classification tasks, and should a new task/class be needed, retraining of the models would be required. To tackle this challenge, our work explores the use of Large Language Models (LLMs) to perform traditional computational tasks, an effort toward eliminating the requirement for explicit model training.

To monitor CIFs in a specified AOI, our proposed approach starts by acquiring all critical facilities in the area from Open Street Maps (OSM) through automatic APIs. Social media data collected about an ongoing disaster is processed through an LLM to generate embeddings and stored in a vector database.[1] Embedding of a message encapsulates its semantics and thus yields better search results. Therefore, for each CIF obtained from OSM, we query the vector database, where raw embeddings are stored, to retrieve messages pertinent to the specified CIF. The retrieved messages are then analyzed by LLMs to identify *(i)* the impacts reported in each message, *(ii)* the severity of the reported impacts, and *(iii)* the operational status of the CIF. We perform all the tasks in a zero-shot setting and evaluate each step of the proposed methodology using standard evaluation metrics.

Through extensive experimentation, we show both the strengths and limitations of LLMs in handling various computational tasks associated with the processing of social media messages. Notably, we showcase diverse configurations of our CIF retrieval query and the corresponding advantages they offer. For instance, our CIF query consisting of a fixed general term (e.g., "disaster impacts") yields better outcomes compared with a detailed listing of impacts. Furthermore, we note better classification performance of LLMs for individual messages when compared to the classification and inference from a set of messages. Overall, we demonstrate that LLMs, even in the zero-shot setting, hold great potential to replace traditional supervised models. However, there are certain weaknesses as well. Notably, LLMs may struggle with context understanding and misinterpret nuanced or ambiguous language commonly found in social media. Additionally, their performance can be influenced by factors like data biases and the nuances of prompt engineering. Acknowledging these limitations is crucial to refining the application of LLMs in processing social media messages effectively.

The rest of the paper is structured as follows. The following section reviews Related Work. In the Methodology section, we provide details of our proposed approach. Results are presented in the Results section and further deliberated upon in the Discussion section. The paper concludes in the final section.

RELATED WORK

Critical infrastructures are at the heart of economic and social welfare in all countries. Therefore, understanding their risk and resilience to various hazards and disasters have been crucially studied in literature (Yusta et al., 2011). For example, Fekete et al. (2015) explored the use of Geographic Information Systems (GIS) and Remote Sensing (RS) whereas Jovanovic´ et al. (2018) presented an indicator-based approach to assess resilience of critical infrastructures to disasters. To this end, Baloye and Palamuleni (2016) proposed a critical infrastructure-driven spatial database for proactive disaster management. Taking a step further, Qiang (2019) provided a comprehensive assessment of flood exposure of all major critical infrastructures in the United States.

In the last decade, with the advancement of social networks, researchers started to explore the use of real-time social media data during disasters to assess risk and resilience of critical infrastructures. For instance, Mittelsta¨dt et al. (2015) integrated social media data with mobile in-situ data to monitor the disaster impact on critical infrastructure. Later Fan and Mostafavi (2019) developed a graph-based method for detecting credible situation information related to infrastructure disruptions during natural disasters. In another study, Roy et al. (2020) presented a multilabel classification approach to identify hurricane-induced infrastructure disruptions through a sentiment analysis of social media data. They utilized several supervised machine learning models to classify disruptions in social media messages. Alternatively, Heglund et al. (2021) demonstrated the utility of social media data for critical infrastructure resilience based on statistical analysis and forecasting methods. While plethora of these methods utilize conventional pipelines of message filtration, trend detection, and other case-specific algorithms, our methodology leverages more generalized inferential capabilities of the large language models (LLMs).

[1]The list of CIFs, data, and embeddings are available upon request.

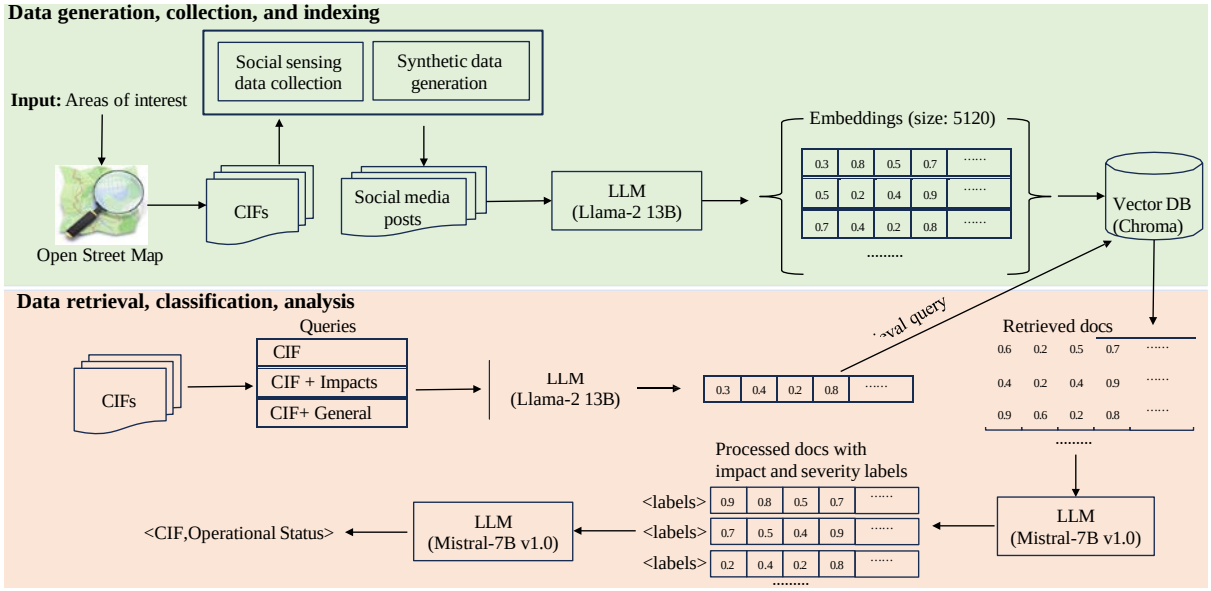


Figure 1. High-level methodology detailing two pipelines: *(i)* data generation, collection, and indexing, *(ii)* data retrieval, classification, and analysis

With the emergence of pre-trained LLMs such as GPT3.5 (Brown et al., 2020), GPT4 (OpenAI et al., 2023), Llama2 (Touvron et al., 2023), Mistral 7B (Jiang et al., 2023), more versatile capabilities have become available for research. For instance, Colverd et al. (2023) has recently studied the use of LLMs for generating flood disaster impact reports via a system initiated by a search query phrase, which is then utilized to gather textual data about the disaster from the web. Another recent use of LLM is geotagging of tweets where few-shot learning is utilized for identification of various location descriptors in tweets such as door number addresses, road segments, and road intersections (Hu et al., 2023).

Inspired by this trend, we capitalize on an LLM for real-time monitoring of critical infrastructure facilities (CIFs) during disasters by classifying their situations in three crucial aspects including the type of impact, its severity, and the operational status. Additionally, to compensate for lack of a publicly available dataset with situational information about CIFs, we again revert to an LLM for its generative capabilities to create synthetic data to test our CIF monitoring pipeline.

METHODOLOGY

Overview

The primary objective of this work is to perform real-time monitoring of various Critical Infrastructure Facilities (CIFs) amidst large-scale disaster scenarios. Specifically, we aim to obtain three types of updates for each CIF in the given area of interest, including *(i)* identifying the disaster impact (e.g., damaged, burned) to the CIF, *(ii)* assessing the severity of the impact (e.g., moderate, severe) and *(iii)* determining the operational status of the facility (e.g., closed, partially closed, or open). We employ social sensing data from social media platforms, harnessing its capability to potentially deliver immediate updates, potentially from eyewitnesses or people possessing any relevant information.

Figure 1 depicts the high-level methodology comprising two pipelines: *(i)* data collection and indexing, and *(ii)* data retrieval and processing. The data collection and indexing pipeline performs tasks related to data acquisition, synthetic data generation, embedding generation, and indexing. The pipeline for data retrieval and processing focuses on retrieving pertinent data from the embedding database. Subsequently, the retrieved data are analyzed through LLMs to identify the impact, severity, and operational status of CIFs. Next, we provide detailed descriptions of the important processes of the methodology.

Obtaining CIFs

We utilize Nominatim APIs2 from Open Street Map (OSM) (OpenStreetMap contributors, 2017) to retrieve information about CIFs within a specified geographical area of interest (AOI). In this work, we focus on two regions

2https://nominatim.openstreetmap.org/

prone to natural disasters: Christchurch, New Zealand, susceptible to earthquakes, and Broward County, Florida, USA, susceptible to hurricanes. We note that for both AOIs we also have Twitter data previously collected from these locations during past disasters. To extract CIF data for these AOIs, we construct queries that include the AOI name and the type of infrastructure, including hospitals, fire stations, rail stations, educational facilities, bridges, roads, and tunnels. Example queries include "Christchurch hospitals," "Christchurch fire stations," etc. Employing this approach, we gathered 58 CIFs from Christchurch and 82 CIFs from Broward County, along with associated metadata such as names, addresses, and geographic coordinates.

Data Curation: Real and Synthetic

Social media data about CIFs is central to our study. However, to the best of our knowledge, there is no publicly available social media data containing tagged CIFs and their associated impacts, severity, and operational statuses. To address this gap, we leverage generative AI to create synthetic data (i.e., tweets) that include real-world CIFs from our selected AOIs. In particular, we utilize the open-source Llama-2 13B model (Touvron et al., 2023) to generate tweets reporting diverse impact scenarios on CIFs retrieved from OSM. Moreover, for each generated tweet, we ask the model to provide the reported impact along with its corresponding severity labels. The specific prompt employed for tweet generation for Broward County is as follows:

> *Generate 15 diverse tweets describing the impact of a Category-5 hurricane on [CIF NAME & ADDRESS]. The disaster triggered sub-events such as tornado, storm surge, burst of rain, strong winds, resulting in varied impacts like flooded, collapsed, submerged, damaged, destroyed, cracked, etc.*
>
> *Ensure linguistic diversity in each tweet, providing unique insights into the impact and its severity. In some tweets, include the infrastructure's address. Aim for tweet lengths between 100 to 250 characters, and avoid using emojis. Tag each tweet with the type of impact (e.g., damaged, destroyed) and its severity (e.g., low, mild, severe). These tags will be used for training classifiers.*
>
> *Always include 2 tags at the 'end' of the generated tweet with the following template: (Tags: ***** , *****)*

For instance, the response returned by the LLM for CIF *"Aventura Hospital and Medical Center & 20900, Biscayne Boulevard, Aventura, Miami-Dade County, Florida, 33180, United States"* looks as follows (showing first three tweets for brevity):

> *Aventura Hospital's emergency room is flooded after a burst of rain caused by the Category 5 hurricane. Patients and staff are being evacuated to safer areas. (Tags: Flooded, Mild)*
>
> *Strong winds from the hurricane have caused significant damage to Aventura Hospital's roof and windows, leaving many areas closed off to patients and staff. (Tags: Damaged, Severe)*
>
> *The storm surge from the hurricane has submerged the lower levels of Aventura Hospital, forcing patients and staff to be evacuated to higher floors. (Tags: Submerged, Severe)*

A similar prompt is used for Christchurch where the first paragraph is adapted for an earthquake scenario as follows:

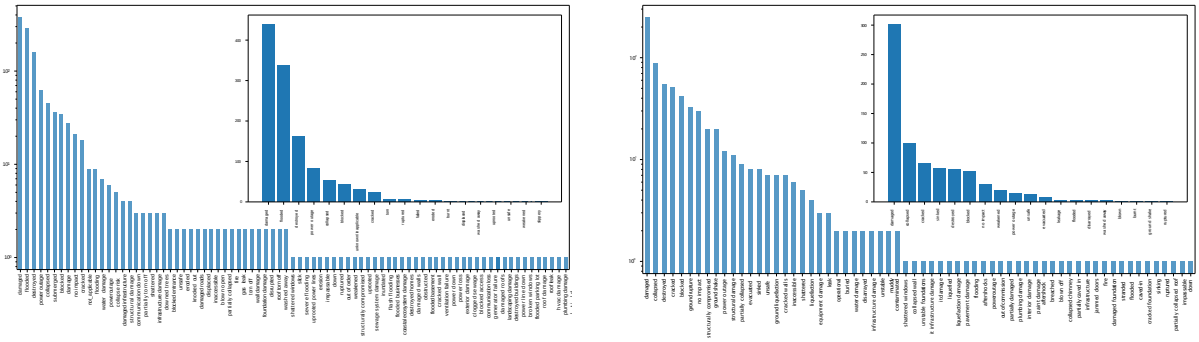
> *Generate 15 diverse tweets describing the impact of a severe earthquake on Christchruch's [CIF NAME & ADDRESS]. The disaster triggered sub-events such as ground shake, landslide, liquefaction, ground rupture, aftershock, resulting in varied impacts like collapsed, cracked, damaged, destroyed, etc.*

Subsequently, we utilize the same LLM employed for tweet generation to assign an operational status label to each generated tweet. The model is restricted to select one of four operational statuses: "open," "closed," "partially open," "partially closed," and "unknown" if no operational status is explicitly mentioned in the tweet. The prompt to obtain the operational status is as follows:

> *Your task is to analyze the provided tweet and determine the operational status of the mentioned infrastructure. The operational status could include descriptors such as open, closed, partially open, partially closed, or unknown.*
>
> *Tweet: [TWEET]*
>
> *Operational status:*

A sample response from the above prompt for the tweet *"The computer lab at Christchurch Girls' High School has been closed due to a collapsed ceiling. Students are using temporary facilities until further notice. #ChristchurchEarthquake"* is as follows:

> *Operational status: closed*



(a) Broward County (b) Christchurch

Figure 2. Distribution of impact labels in the synthetic data for (a) Broward County and (b) Christchurch. The outer charts with light blue bars correspond to the LLM-generated raw tags whereas the overlaid smaller charts with dark blue bars show the manually pruned ground-truth tags.

In total, we generated 728 tweets for Christchurch for 58 CIFs and 1,205 tweets for Broward County against 82 CIFs. Although instructed, the model does not always generate 15 tweets for each CIF. Tweets with incorrect tags/labels were manually corrected. As we did not control the prompt to enforce a closed-set of impact types, the resulting list of impact tags is open-ended (N=83), which also allowed for more diversity in the generated tweets. Figure 2 shows the distribution of the resulting tags for each AOI. We noticed multiple similar impact labels (e.g., flooded, flooding, inundated, etc.) in the model responses, which we merged into a unified list of tags, as illustrated in Table 1. The sub-figures in Figure 2 depict the distribution of the refined tag list. Note that, unlike impact tags, we utilized a closed-set of tags for classification of severity (i.e., severe, moderate, mild, and unknown) and operational status (i.e., open, closed, partially open, partially closed, and unknown).

Later, for each CIF, we dispersed its generated tweets across four six-hour time intervals (spanning a 24-hour time frame) by considering their impact order (e.g., "destroyed" followed by "damaged" or "cracked" and so on). The impact ordering ensures a coherent progression of the impact to a CIF. This step is crucial for simulating a real-world setting in our subsequent experiments. Moreover, given that several past studies highlight the prevalence of noise and the limited relevant information in real-world tweets (Grace, 2021; Kumar et al., 2017), we aim to ensure the realism of our dataset. To achieve this, we introduce noise into our generated data with a signal-to-noise ratio of 2% (signal) to 98% (noise). For this purpose, we incorporated tweets collected during past disasters for both AOIs. Specifically, for Christchurch, we used tweets collected after a magnitude-7.1 earthquake struck New Zealand on September 2nd, 2016. For Broward County, we utilized tweets collected after Hurricane Ian hit Florida in September 2022. In total, we prepared 36,286 and 60,062 tweets (including both synthetic and real) for Christchurch and Broward County, respectively. To ensure that CIFs also appear in non-impact tweets, we injected CIF names into 8% of the noisy tweets in each time interval.

Finally, to assign an overall status for a CIF within each time interval, we deduce its operational status. To achieve this, we chronologically order tweets for a CIF and use the operational status of the last tweet as an overall status. If the operational status of the last tweet is "unknown", we use the last valid status.

Generating and Indexing Tweet Embeddings

Next, we employ Llama-2 13B (Touvron et al., 2023) to compute embeddings for all the tweets, including synthetic and real. Each tweet is converted into a 5120-dimensional embedding vector which is the default embedding size of Llama-2 13B. This embedding is then indexed into ChromaDB[3], a well-known open-source vector database. These embeddings serve as the basis for semantic search, enabling the retrieval of tweets specifically mentioning certain CIFs.

Retrieval of CIF Tweets

The first step in determining the impact and operational status of a CIF involves retrieving tweets referencing a specific CIF. To achieve this, we explore three types of retrieval queries: *(i)* a query comprising only the CIF name, *(ii)* a query involving both the CIF name and a set of impact terms such as damaged, cracked, burned, destroyed, submerged, and *(iii)* a query combining the CIF name with the term "disaster impacts." For each query, we first

3https://www.trychroma.com/

Table 1. Raw impact labels and their mapping to final impact taxonomy for LLM-based classification

Consolidated labels	Raw model labels
damaged	aftershock damage, damage, damaged, broken windows, damaged infrastructure, damaged roads, damaged roofs, damaged walls, electrical damage, coastal ecosystem damage, foundation damage, landscaping damage, exterior damage, hvac damage, infrastructure damage, roof damage, plumbing damage, sewage system damage, structural damage, wall damage, water damage, equipment damage, damaged foundation, interior damage, IT damage, liquefaction damage, partially damaged, pavement damage, plumbing damage, IT infrastructure damage, equipment damage, paint damage
flooded	flooded basement, flooded businesses, flooded parking lot, flooding, severe flooding, flash flooding, submerged, inundated
destroyed	destroyed, destroyed buildings, destroyed homes, rubble
weakened	weakened, structurally compromised
cracked	cracked, shattered, shattered windows, debris, cracked wall, cracked walls
blocked	blocked access, blocked entrance, impassable, obstructed, clogged sewage, inaccessible, partially blocked, landslide, jammed doors
torn	partially torn off, roof torn off
power outage	power line down, power loss, power outage, power down, communication down, communication loss, down, knocked out, disrupted, downed trees, offline
ruptured	blown open, gas leak, roof leak
collapsed	partially collapsed, collapsed, crushed, collapse risk, collapsed wall, collapsed chimney, partially collapsed roof
failed	generator failure, ventilation failure, out of order
uprooted	uprooted, uprooted power lines
eroded	eroded, erosion
washed away	washed away, muddy
slippery	slippery, slick
displaced	displaced
blown	blown, blown off
burnt	burnt, burning, fire
unsafe	unsafe, unstable, structurally compromised, uninhabitable, contaminated
leakage	leak, gas leak, gasleak
sinked	ground liquefaction, liquefaction, buried, sinked, sinking, caved-in, ground rupture, liquefied
unknown/inapplicable	not applicable, not humanitarian, no impact, unknown

obtain its embeddings from the Llama-2 13B model and subsequently retrieve the top-K (ranging from 5 to 50 with a step size of 5) tweets by querying ChromaDB using cosine similarity.

Top-K results for each query are then used to compute Average Precision (AP) using Equation (1).

$$\text{AP@}K = \frac{1}{\text{GTP@}K} \sum_{\substack{k=5n \\ 1 \le n \le N}} \text{P@}k \times \text{rel@}k, \text{ for } K = 5N \text{ and } N = \{1, \ldots, 10\} \tag{1}$$

where

$$\text{P@}k = \frac{\text{Number of relevant documents retrieved @}k}{k}$$

$$\text{rel@}k = \begin{cases} 1, & \text{if } k^{th} \text{ document is relevant} \\ 0, & \text{otherwise} \end{cases}$$

$$\text{GTP@}K = \min(K, \text{Total number of relevant documents})$$

In Equation (1), we normalize the total sum by the number of 'retrievable' ground-truth documents or *signals* (i.e., Ground Truth Positive or GTP). Finally, to determine the overall retrieval performance of a given query, we

compute mean Average Precision (i.e., mAP) in a time interval $t \in$ {0h-6h, 6h-12h, 12h-18h, 18h-24h, 0-24h} using Equation (2).

$$\text{mAP@}K^t = \frac{1}{M}\sum_{i=1}^{M} \text{AP@}K_i^t \text{, for } K = 5N \text{ and } N = \{1, \ldots, 10\} \quad (2)$$

Tweet Classification for Impact, Severity, and Operational Status

We employ the Mistral 7B v1.0 model (Jiang et al., 2023) for classifying the retrieved tweets that are identified as relevant by the CIF queries. These tweets are likely to encompass information about impacts experienced by critical facilities. To perform the impact and severity classification of each tweet, we use the following prompt template:

> *Your task is to analyze the provided tweet and determine the impacts of the disaster on the mentioned infrastructure. Please include only impact descriptors from the list such as blocked, blown, buried, burnt, collapsed, cracked, damaged, destroyed, displaced, disrupted, eroded, failed, flooded, ground liquefaction, ground shake, leakage, muddy, power outage, ruptured, slippery, torn, unsafe, uprooted, washed away, weakened or not applicable, and severity such as severe, mild, moderate, unknown.*
>
> *Tweet: [TWEET]*
>
> *Infrastructure impact:*
> *Infrastructure severity:*

The response returned by the model to an example tweet such as *"The Christchurch Public Hospital's radiology equipment is malfunctioning due to the earthquake, making it difficult to diagnose patients. #ChristchurchEarthquake"* looks as follows:

> *Infrastructure impact: damaged*
> *Infrastructure severity: moderate*

Finally, tweets that are tagged with an impact type are further analyzed to infer the operational status of CIFs. For this purpose, we use the Mistral 7B v1.0 model with the following prompt:

> *Your task is to analyze the provided tweet and determine the operational status of the mentioned infrastructure. The operational status could include descriptors such as open, closed, partially open, partially closed, or unknown.*
>
> *Tweet: [TWEET]*
> *Operational status:*

The response returned by Mistral 7B v1.0 model to an example tweet such as *"Ground shake from Christchurch earthquake caused significant damageto Wigram Fire Station's foundation, rendering it unstable. #ChristchurchEarthquake"* looks as follows:

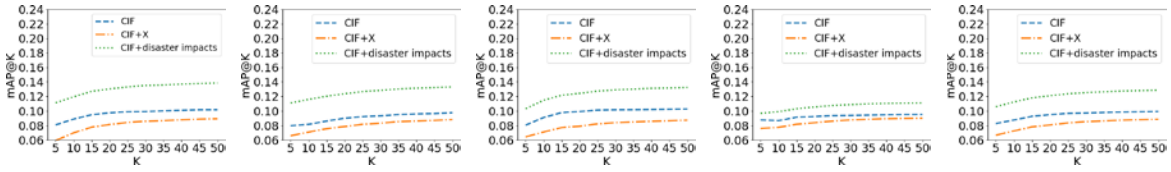
> *Operational status: closed*

To evaluate the performance of impact, severity, and operational status classification, we compare the labels predicted by the Mistral 7B model with the ground-truth labels obtained from the Llama-2 model and report results using standard evaluation metrics such as precision, recall, and F1-score.

Obtaining Overall Operational Status of CIFs

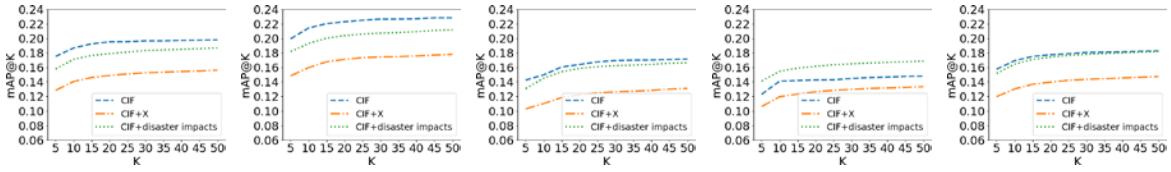
The last step in the tweet classification and analysis pipeline involves extracting the overall operational status of tweets retrieved through a CIF query. The set of tweets resulting from a CIF query may encompass information about the queried CIF, but these tweets may vary in terms of containing impact-related details. Therefore, to determine the overall operational status of a CIF at a specific time, we only consider tweets containing impacts (relying on the impact labels obtained in the previous step). To achieve this, we employ the Mistral 7B v1.0 model with the following prompt to retrieve the overall status.

> *Your task is to analyze the tweets given below and deduce the operational status of a facility, named [CIF]. Since these tweets are retrieved based on the facility name, it's possible that some tweets may not pertain to the given facility. Focus solely on the tweets pertinent to [CIF] and derive the most recent operational status for the facility. Your operational status label must be one of these: open, closed, partially open, partially closed, or unknown.*
>
> *Tweet: [TWEET 1]*
> *Tweet: [TWEET 2]*
> *Tweet: [TWEET 3]*
> *...*
>
> *operational_status:*



(a) 0h-6h (b) 6h-12h (c) 12h-18h (d) 18h-24h (e) 0h-24h

Figure 3. Performance comparison of retrieval queries for Broward County



(a) 0h-6h (b) 6h-12h (c) 12h-18h (d) 18h-24h (e) 0h-24h

Figure 4. Performance comparison of retrieval queries for Christchurch

RESULTS

In this section we present our results for the retrieval and classification tasks described in the Methodology section.

Retrieval of CIF Tweets

For each AOI, we considered three types of queries to fetch relevant tweets pertaining to each CIF:

(i) *CIF* — a query comprising only the CIF name

(ii) *CIF + X* — a query involving both the CIF name and a set of impact terms such as damaged, cracked, burned, destroyed, submerged, etc. (denoted as X for brevity)

(iii) *CIF + disaster impacts* — a query combining the CIF name with the term "disaster impacts"

Specifically, for Broward County *X* comprises *flooded, submerged, damaged, destroyed, weakened, cracked, blocked, torn, power outage, ruptured, collapsed, failed, uprooted, eroded, burnt, washed away, slippery, displaced, disrupted* whereas for Christchurch it consists of *flooded, destroyed, leak, blocked, cracked, ground liquefaction, power outage, ruptured, buried, collapsed, ground shake, unsafe, muddy*.

We also experimented with K to assess its effect on the retrieval performance. That is, we ran queries for top-K where we changed K from 5 to 50 with a step size of 5. We then computed the overall mAP@K in each time interval using Equation (2) to evaluate the overall performance of each query type. Figures 3 and 4 show the performance plots for all query types across different K values and time intervals for Broward County and Christchurch, respectively. The plots initially show an increasing trend in performance with increasing K but the performance improvement generally slows down after $K = 30$. For Broward County, *"CIF + disaster impacts"* queries outperform the other queries by a big margin whereas for Christchurch *"CIF"* queries hold a slight edge over the other queries in the six-hour time intervals but perform only on par with *"CIF + disaster impacts"* in the 0h-24h time interval.

Tables 2 and 3 summarize the mAP@K scores for $K = 50$ for Broward County and Christchurch, respectively. Again we observe that the scores for the *"CIF + disaster impacts"* queries are highest in all time intervals for Broward County whereas *"CIF"* and *"CIF + disaster impacts"* queries perform on par for Christchurch. However, it is important to note that, for both AOIs, *"CIF + disaster impacts"* queries retrieve the highest number of relevant tweets among all query types in 0h-24h time interval, i.e., 534 out of 1205 relevant tweets in Broward County and 402 out of 728 relevant tweets in Christchurch. Therefore, we conclude that *"CIF + disaster impacts"* query type yields the best overall retrieval performance, and hence, we use it as the default query type in the proposed pipeline.

We note that the resulting mAP scores are generally low and the effective hit rate (i.e., ratio of relevant retrieved tweets to total number of retrieved tweets) is only slightly above 3% for both AOIs. To investigate this further, we color coded and plotted the distribution of relevant and irrelevant tweets in Figure 5. We additionally sub-divided irrelevant tweets belonging to either noise or other CIF that we generated. From Figure 5, we observe that although a large portion of the distribution for each CIF contains irrelevant tweets, a significant portion of irrelevant tweets is

Table 2. Retrieval performance (mAP@50) of the Broward County queries

	Time Intervals					
Queries	0h–6h	6h–12h	12h–18h	18h–24h	0h–24h	Y
CIF	0.102	0.098	0.103	0.095	0.099	303
CIF + X	0.089	0.088	0.087	0.090	0.089	484
CIF + "disaster impacts"	**0.139**	**0.133**	**0.133**	**0.111**	**0.129**	**534**

X = flooded, submerged, damaged, destroyed, weakened, cracked, blocked, torn, power outage, ruptured, collapsed, failed, uprooted, eroded, burnt, washed away, slippery, displaced, disrupted
Y = Number of relevant tweets retrieved out of 1205 in 0-24h time interval

Table 3. Retrieval performance (mAP@50) of the Christchurch queries

	Time Intervals					
Queries	0h–6h	6h–12h	12h–18h	18h–24h	0h–24h	Y
CIF	**0.198**	**0.229**	**0.172**	**0.148**	**0.183**	288
CIF + X	0.156	0.178	0.131	0.134	0.148	326
CIF + "disaster impacts"	0.187	0.212	0.167	0.169	0.182	**402**

X = flooded, destroyed, leak, blocked, cracked, ground liquefaction, power outage, ruptured, buried, collapsed, ground shake, unsafe, muddy
Y = Number of relevant tweets retrieved out of 728 in 0-24h time interval

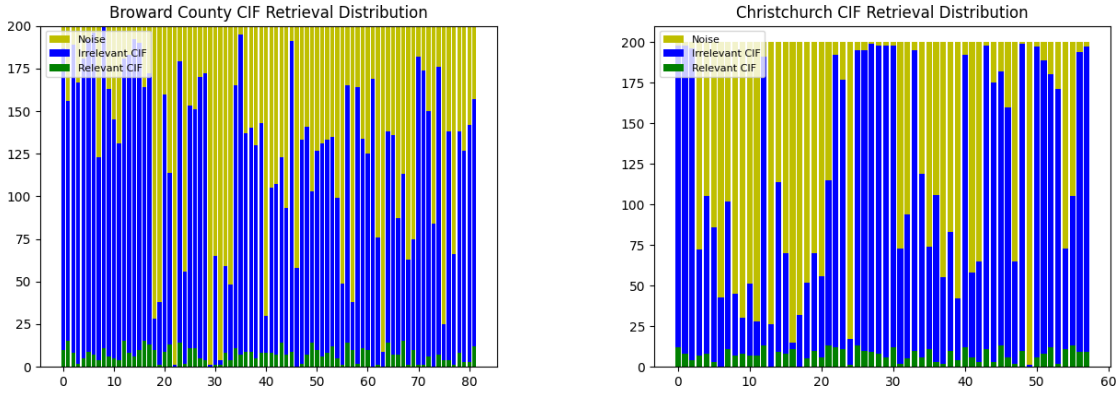
occupied by tweets describing impacts to other CIFs and share some semantic similarity with the relevant tweet. For each query on average roughly 60% of the retrieved results belong to a certain CIF (not necessarily the correct one) and the remaining contains pure noise. Although achieving a higher hit rate is always desirable, we defer further improvements for the retrieval process to our future work.

Tweet Classification for Impact, Severity, and Operational Status

Here we compare the ground-truth impact, severity, and operational status labels obtained from the Llama-2 13B model in the dataset generation stage with the labels inferred by the Mistral 7B v1.0 model in the data analysis and classification stage. Because we do not expect the classification models to depend on a specific CIF or time interval, we evaluate the classification performance over all tweets across all CIFs and time intervals using standard evaluation metrics such as Precision, Recall, and F1-score. Table 4 summarizes the results of tweet-level impact, severity, and operational status classification for both AOIs. The results indicate that the proposed LLM-based classification models achieve reasonable performance for the task at hand. In general, the models achieve better precision than recall in all three tasks. However, the differences between precision and recall are bigger in the operational status classification task, which leads to the lowest F1-scores among all the tasks.

To take a closer look at the impact classification performance on the synthetic data, we provided confusion matrices in Figure 6. For Broward County, the most prominent impact classes such as *damaged* and *flooded* are correctly classified most of the time, but it is notable that around one third of the *flooded* tweets are classified as *damaged* (N=112). This is understandable because *damaged* class has a broader meaning and coverage that overlaps with many other impact types including *flooded*. Similarly, *blocked* (N=20), *destroyed* (N=56), and *power outage* (N=58) tweets are usually classified as *damaged*. However, the opposite is not correct, and in line with this expectation, we see much fewer *damaged* tweets getting labeled as *flooded* (N=2), *blocked* (N=13), *destroyed* (N=2), and *power outage* (N=1). We also observe that two thirds of the *unknown/inapplicable* tweets (N=20) are confused as *damaged*, which is not very desirable and needs closer investigation. For Christchurch, we observe similar trends where, as the most prominent impact class, *damaged* tweets are correctly classified most of the time (N=302), but many other impact tweets such as *blocked* (N=17), *collapsed* (N=51), *cracked* (N=55), *destroyed* (N=15), and *weakened* (N=15) are also tagged as *damaged*. Moreover, majority of the *unknown/inapplicable* tweets (N=23) are confused as *damaged* tweets. Altogether these observations suggest that it is challenging to get the Mistral 7B v1.0 model (or any other LLM for that matter) distinguish between fine-grained impact types.

We then also analyze the impact classification performance on the output of the retrieval system which contains noisy data due to imperfect nature of the retrieval process. Here we assume that ground-truth labels for the noise tweets belong to *unknown/inapplicable* class for all classification tasks. Table 5 summarizes the scores for impact,



(a) Broward County **(b) Christchurch**

Figure 5. Signal-to-noise distribution of retrieved tweets for each CIF in (a) Broward County and (b) Christchurch

Table 4. Performance scores for impact, severity, and operational status classification on the synthetic data

	Impact			Severity			Operational Status		
AOI	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Broward County	0.660	0.563	0.565	0.668	0.512	0.566	0.805	0.349	0.445
Christchurch	0.603	0.570	0.569	0.550	0.458	0.490	0.655	0.280	0.342

severity, and operational status classification on the retrieved tweets. These scores are higher thanks to the model's ability to correctly identify the (incorrectly retrieved) noise tweets as "unknown/inapplicable."

Figure 7 visualizes confusion matrices for further analysis of the classification results of the retrieved tweets. For both AOIs, the confusion matrices are dominated by the correctly identified noise tweets as *unknown/inapplicable* (i.e., N=4987 for Broward County and N=1178 for Christchurch). Besides, for Broward County, majority of the tweets are correctly classified as *damaged* (N=1536), *destroyed* (N=1336), *flooded* (N=1781), *blocked* (N=381), and *power outage* (N=365). However, many ground-truth *damaged* tweets were also labeled incorrectly as *torn* (N=467) or *power outage* (N=273) and many *flooded* tweets as *damaged* (N=308) or *power outage* (N=285). Whereas for Christchurch, while *damaged* (N=252), *destroyed* (N=90), *collapsed* (N=108), and *blocked* (N=47) tweets are classified correctly in general, we see that most *cracked* tweets were confused as *damaged* (N=84) and *damaged* tweets as *collapsed* (N=34). We also see that impacts such as *burnt, ruptured, uprooted, sink,* and *weakened* are oftentimes misclassified. Another interesting observation is that so many tweets are classified (incorrectly more than half of the time) as *power outage* in Broward County whereas almost no tweet gets classified as *power outage* in Christchurch.

Overall Operational Status Classification of CIFs

Tweets retrieved with the impact labels were processed using Mistral 7B v1.0 to infer the overall operational status of a target CIF. The results are provided in Table 6. While exhibiting a reasonable precision, the model struggled with the recall. This could be due to the retrieval process, which incorporates messages from other CIFs and noise. Even though we mitigated the noise by focusing solely on impact-related messages, messages about other pertinent CIFs still pose a challenge for the model.

DISCUSSION

Besides illustrating the opportunities LLMs have to offer for real-time monitoring of CIFs by analyzing social media data during disasters, this study also reveals various challenges to be tackled along the way. For instance, an important challenge is how to instruct LLMs to generate a realistic timeline of short messages (i.e., tweets) about CIFs during disasters that are coherent and consistent both spatially and temporally. Another important challenge is how to guide LLMs to eliminate irrelevant messages successfully in the retrieval process. Yet another important challenge is how to prompt LLMs in a zero-shot setting to discern fine-grained details about disaster impact types and severity affecting CIFs as well as to infer the operational status of CIFs from a set of impact and severity reports.

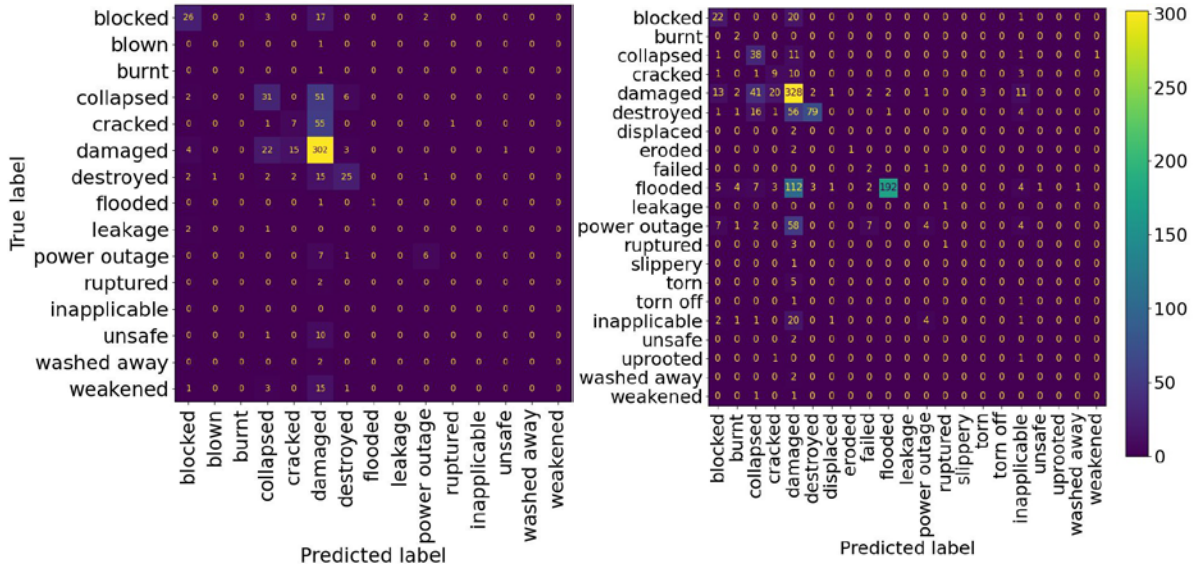


Figure 6. Confusion matrices for all signal impact classification (left) Broward County and (right) Christchurch

Table 5. Performance scores for impact, severity, and operational status classification on the retrieved tweets

	Impact			Severity			Operational Status		
AOI	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Broward County	0.821	0.747	0.769	0.687	0.6445	0.6442	0.517	0.352	0.399
Christchurch	0.841	0.805	0.777	0.817	0.794	0.797	0.919	0.779	0.809

On the data side, we opted for open-ended generation of impact types. This allowed us to generate a quite diverse set of fine-grained impact types (N=83), a property sought by stakeholders to make key decisions. However, this strategy also resulted in a long-tail distribution of impact types where the generic *damaged* became the most prominent impact type by a large margin while more than two thirds of the impact types appeared only once or twice in the entire dataset (see Figure 2). An alternative strategy could be to impose more control over the distribution and ask LLM to generate tweets for a specific impact type at a time (potentially at the expense of reduced diversity).

On the retrieval side, our current approach experienced difficulty in eliminating the irrelevant content for some CIFs more than others (see Figure 5). This could be attributed partly confusing nature of similar CIF or partial/shortened CIF names and partly to weaknesses of the retrieval method employed. In future studies, we plan to explore more sophisticated retrieval techniques.

On the classification side, we observed weaknesses of LLMs in distinguishing granular details about disaster impact types as analyzed through confusion matrices in Figures 6 and 7. There is also the fact that there are multiple correct answers many times and formulating the problem as a single-label classification becomes problematic as illustrated by the examples below:

> *Case 1: The Category 5 hurricane has caused significant damage to the South Area Alternative School's electrical system, leaving the building without power. The school's future is uncertain.*
>
> *Llama-2 13B (ground truth): Power outage*
> *Mistral 7B v1.0 (classified): Damaged*
>
> *Case 2: The hurricane has caused a tree to fall on top of a building near Broward Health Imperial Point, causing significant damage and blocking access to the hospital.*
>
> *Llama-2 13B (ground truth): Damaged*
> *Mistral 7B v1.0 (classified): Blocked*
>
> *Case 3: The University of Florida Field Laboratory's interior is flooded, causing significant damage to research equipment and infrastructure. The facility is closed indefinitely.*
>
> *Llama-2 13B (ground truth): Flooded*
> *Mistral 7B v1.0 (classified): Damaged*

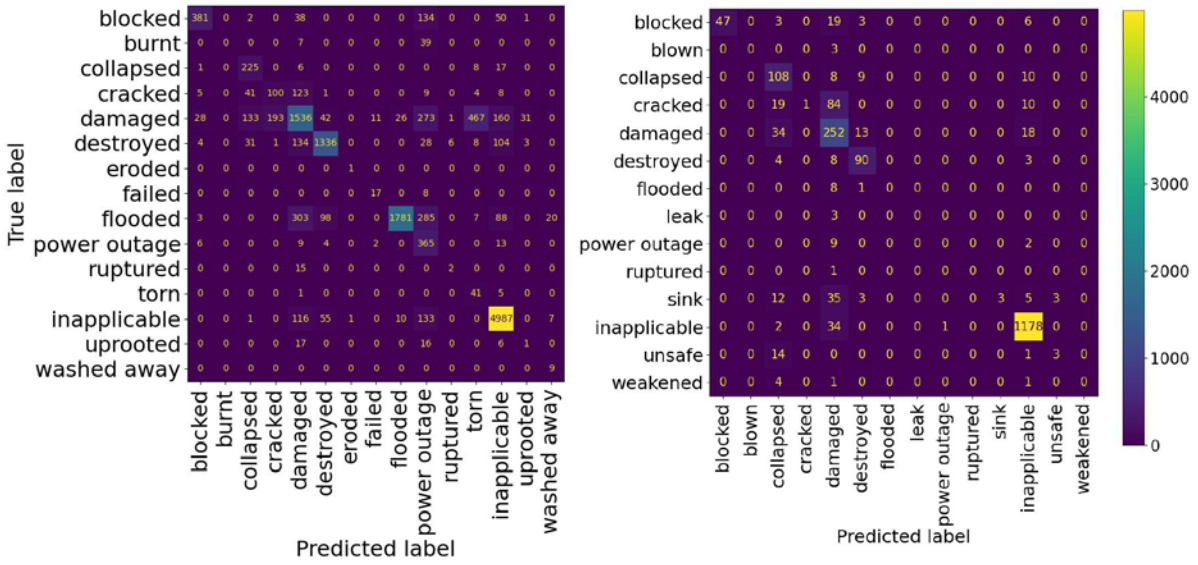


Figure 7. Confusion matrices for impact classification of retrieved tweets (left) Broward County and (right) Christchurch

Table 6. Performance scores for overall operational status classification on the retrieved results

	Overall Operational Status		
AOI	Precision	Recall	F1-score
Broward County	0.522	0.305	0.216
Christchurch	0.470	0.248	0.197

Case 4: Honey Hill Fire Station's front entrance crushed by fallen tree. Emergency responders working to clear debris and reach those trapped.

Llama-2 13B (ground truth): Collapsed
Mistral 7B v1.0 (classified): Blocked

In all cases, both ground-truth and classified labels seem reasonable with subtle differences stemming from how each LLM interprets the overall situation.

Among all the tasks performed by LLMs, the most challenging proved to be inferring the overall operational status of a CIF. Given that the retrieval process may include messages that are semantically akin to a CIF query, potentially encompassing both, messages about other CIFs and noise, the model tends to face confusion when deducing the overall operational status. We remark that enhancing the retrieval process will significantly improve the performance of the inference task.

LIMITATIONS AND FUTURE WORK

In this section, we describe the biases and limitations of this study. Our main source of bias comes from the dataset utilized in this study. Due to the unavailability of real-world curated datasets describing impacts to CIFs, we utilized Llama-2 13B to generate the relevant tweets as well as the matching labels. In Figure 2, we note that the distributions of impacts generated by Llama-2 13B mainly cover broader impacts such as *damaged, destroyed, flooded and collapsed*. Although Llama-2 13B also generated fine-grained impacts such as *shattered windows, ruptured pipe, HVAC damage and muddy road etc*, such distribution might not be representative of the distribution of impacts in real-world disaster. Likewise, the class/label imbalance in the dataset also limits us from fully accessing the capabilities of Mistral 7B v1.0 for the classification of various low-frequency classes. Moreover, the progression of impacts to CIFs in 24-hour time window was done by manually assigning the precedence of impacts to a CIF by the authors. This in part is also a bias that could have affected the prediction of the overall operational status performed in this study.

Among the limitations, the most predominant ones are the language usage, the choice of LLM models and the nature of disasters selected for this study. Unlike real-world tweets, the synthetic tweets generated by Llama-2 13B

were free from any typos, slang or grammatical errors. This study, therefore, does not fully assess the sensitivity of LLM-based classification to such inconsistencies. Moreover, we limited the type of disasters as well the choice of AOI used in this study to only two. This significantly limits the vocabulary of impacts as well as the nature of the content of the tweets in the generated dataset. The choice of AOI used, i.e., Christchurch and Broward County, also resulted in the generation and classification of only English medium tweets.

Likewise, this study is also limited to one particular LLM for the data generation i.e., Llama-2 13B and another LLM for the classification task i.e., Mistral 7B v1.0. Therefore, further analysis of the quality of data generation and classification of crisis-relevant tweets by other LLMs is deferred to our future work.

CONCLUSION

In this study, we investigated the application of Large Language Models to monitor the operational status of critical facilities in an area during large-scale disasters. We utilized social sensing data (both synthetic and real-world) from X (formerly Twitter) for two distinct locations and performed their analysis using LLMs for various computational tasks, including retrieval, classification, and inference. The outcomes are presented using standard evaluation metrics and further discussed to highlight the strengths and weaknesses of LLMs. In summary, we emphasize that the use of LLMs in the zero-shot setting, as compared to supervised models requiring training data, demonstrates reasonable performance. However, challenges arise for LLMs when dealing with complex contextual information provided through lengthy prompts. Additionally, we outline various future directions that we anticipate to be beneficial for the community to build upon.

REFERENCES

Alam, F., Qazi, U., Imran, M., & Ofli, F. (2021). Humaid: Human-annotated disaster incidents data from twitter with deep learning benchmarks. *Proceedings of the International AAAI Conference on Web and social media*, *15*, 933–942.

Auerswald, P., Branscomb, L. M., La Porte, T. M., Michel-Kerjan, E., & MICHEL-KERJAN, E. (2005). The challenge of protecting critical infrastructure. *Issues in Science and Technology*, *22*(1), 77–83.

Baloye, D. O., & Palamuleni, L. G. (2016). Modelling a critical infrastructure-driven spatial database for proactive disaster management: A developing country context. *Ja`mba´: Journal of Disaster Risk Studies*, *8*(1), 1–14.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877–1901.

Castillo, C. (2016). *Big crisis data: Social media in disasters and time-critical situations*. Cambridge University Press.

Colverd, G., Darm, P., Silverberg, L., & Kasmanoff, N. (2023). Floodbrain: Flood disaster reporting by web-based retrieval augmented generation with an llm. *NeurIPS Workshop on Artificial Intelligence for Humanitarian Assistance and Disaster Response*.

Fan, C., & Mostafavi, A. (2019). A graph-based method for social sensing of infrastructure disruptions in disasters. *Computer-Aided Civil and Infrastructure Engineering*, *34*(12), 1055–1070.

Fekete, A., Tzavella, K., Armas, I., Binner, J., Garschagen, M., Giupponi, C., Mojtahed, V., Pettita, M., Schneiderbauer, S., & Serre, D. (2015). Critical data source; tool or even infrastructure? challenges of geographic information systems and remote sensing for disaster risk governance. *ISPRS International Journal of Geo-Information*, *4*(4), 1848–1869.

Grace, R. (2021). Toponym usage in social media in emergencies. *International Journal of Disaster Risk Reduction*, *52*, 101923.

Heglund, J., Hopkinson, K. M., & Tran, H. T. (2021). Social sensing: Towards social media as a sensor for resilience in power systems and other critical infrastructures. *Sustainable and Resilient Infrastructure*, *6*(1-2), 94–106.

Hu, Y., Mai, G., Cundy, C., Choi, K., Lao, N., Liu, W., Lakhanpal, G., Zhou, R. Z., & Joseph, K. (2023). Geo-knowledge-guided gpt models improve the extraction of location descriptions from disaster-related social media messages. *International Journal of Geographical Information Science*, *37*(11), 2289–2318.

Imran, M., Castillo, C., Diaz, F., & Vieweg, S. (2015). Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)*, *47*(4), 1–38.

Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D. S., Casas, D. d. l., Bressand, F., Lengyel, G., Lample, G., Saulnier, L., et al. (2023). Mistral 7b. *arXiv preprint arXiv:2310.06825*.

Jovanovic´, A., Øien, K., & Choudhary, A. (2018). An indicator-based approach to assessing resilience of smart critical infrastructures. *Urban Disaster Resilience and Security: Addressing Risks in Societies*, 285–311.

Kumar, A., Singh, J. P., & Rana, N. P. (2017). Authenticity of geo-location and place name in tweets.

Labaka, L., Hernantes, J., & Sarriegi, J. M. (2016). A holistic framework for building critical infrastructure resilience. *Technological Forecasting and Social Change*, *103*, 21–33.

Mittelsta¨dt, S., Wang, X., Eaglin, T., Thom, D., Keim, D., Tolone, W., & Ribarsky, W. (2015). An integrated in-situ approach to impacts from natural disasters on critical infrastructures. *2015 48th Hawaii International Conference on System Sciences*, 1118–1127.

Olteanu, A., Vieweg, S., & Castillo, C. (2015). What to expect when the unexpected happens: Social media communications across crises. *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, 994–1009.

OpenAI, : Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., . . . Zoph, B. (2023). Gpt-4 technical report.

OpenStreetMap contributors. (2017). Planet dump retrieved from https://planet.osm.org.

Pescaroli, G., & Kelman, I. (2017). How critical infrastructure orients international relief in cascading disasters. *Journal of Contingencies and Crisis Management*, *25*(2), 56–67.

Qiang, Y. (2019). Flood exposure of critical infrastructures in the united states. *International Journal of Disaster Risk Reduction*, *39*, 101240.

Qu, Y., Huang, C., Zhang, P., & Zhang, J. (2011). Microblogging after a major disaster in china: A case study of the 2010 yushu earthquake. *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 25–34.

Roy, K. C., Hasan, S., & Mozumder, P. (2020). A multilabel classification approach to identify hurricane-induced infrastructure disruptions using social media data. *Computer-Aided Civil and Infrastructure Engineering*, *35*(12), 1387–1402.

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., et al. (2023). Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Yusta, J. M., Correa, G. J., & Lacal-Ara´ntegui, R. (2011). Methodologies and applications for critical infrastructure protection: State-of-the-art. *Energy policy*, *39*(10), 6100–6119.

Zahra, K., Imran, M., & Ostermann, F. O. (2020). Automatic identification of eyewitness messages on twitter during disasters. *Information processing & management*, *57*(1), 102107.

Empowering Crisis Response Efforts: A Novel Approach to Geolocating Social Media Images for Enhanced Situational Awareness

Hafiz Budi Firmansyah*
Centre Universitaire d'Informatique
University of Geneva, Switzerland
Hafiz.Firmansyah@etu.unige.ch

Jose Luis Fernandez-Marquez
Centre Universitaire d'Informatique
University of Geneva, Switzerland
JoseLuis.Fernandez@unige.ch

Mehmet Oguz Mulayim
Artificial Intelligence Research Institute, IIIA
Spanish National Research Council, CSIC
Cerdanyola del Vallès, Spain

Jorge Gomes
Virtual Operations Support Teams
VOST Europe
Santarém, Portugal

Joao Ribeiro
Virtual Operations Support Teams
VOST Europe
Santarém, Portugal

Valerio Lorini
Joint Research Center
European Commission
Ispra, Italy

ABSTRACT

Social media and online platforms play an important role in assessing the impact of a natural disaster, especially in the immediate aftermath of the event. The information from social media demonstrated the potential to accelerate the response to a crisis. Locating social media images is critically important to help the victims immediately, mobilize community support, and provide enhanced situational awareness. However, finding the location of relevant social media images still challenges humans and computational processes.

This study introduces a social media image classifier aimed at enhancing crowdsourced geolocation. The model is trained using data annotated by experts in disaster risk management. Its main goal is to reduce the difficulty of geolocating images by detecting those that are "easy" to geolocate. We combined the classifier with an ad-hoc crowdsourcing platform and tested it with pictures posted during a crisis. The experimental results indicate that the proposed approach speeds up the geolocation process of social media images while increasing the level of location precision

Keywords

Social media, geolocation, disaster management

INTRODUCTION

Data from social media demonstrates the ability to facilitate quicker response in crisis situations. In 2020, the Joint Research Center European Commission activated a social media analysis to understand the Beirut explosion. Their work proposes a framework to strengthen the use of social media for crisis management that could complement the current Copernicus Mapping service in the initial phases of the crisis. Further developments aim at the automated

*corresponding author

detection of relevant media, their localization and classification providing an initial impact assessment starting from a few moments after the disaster 1. Despite usefullness of social media data in previous example, employing data from social media to evaluate the damage introduces significant challenges. For privacy concerns, major social media platforms remove the meta information about locations by default. Out of millions of posts on Twitter, one of the most used platforms by researchers, only up to 3% are geolocated by users (Huang and Carley, 2020). Furthermore, identifying the location of relevant photos with evidence of damage to organize the first response efficiently presents a challenge. While several techniques have been proposed to automatically geolocate social media content, including images (e.g., Murgese et al., 2022) and text (e.g., Scalia et al., 2022), their predictions need to be more precise to enable their use by practitioners in the field. Combining automatic geolocation with crowdsourcing offers a valuable solution to this problem, with the former providing a preliminary location estimate and the latter being used to enhance the level of precision of that location. The geolocation is then pinpointed on a map to visualize the affected areas. The actual goal of extracting location from social media images is to create a disaster map that could help the first responders locate the events seen on posted images.

Thanks to the combination of Crowdsourcing and Artificial Intelligence methods, this research aims to reduce the overall time of geolocating social media images posted from regions impacted by a crisis and increase the geolocation precision. To pursue the objective, we decrease the difficulty of geolocating images by selecting easy-to-geolocate images. The study leverages past research activities on social media images for the disaster risk management domain to support emergency response as a framework (Rufolo et al., 2021), a platform (Lorini, Panizio, and Castillo, 2022), and a disaster map (Lorini, Rufolo, and Castillo, 2022).

The remainder of the paper is organized as follows: First, we discuss related work and research gaps. Second, we describe our dataset and experiment method based on the research gaps and present the methods we followed to shorten the research gaps and detail the experiments that demonstrate the validity of our approach. Then, we report the experimental results with social media data from four disaster events. Finally, last section concludes this work and provides future research directions.

RELATED WORK

Geolocation prediction on social media involves linking an entity, like images or posts, to a specific location with a high level of accuracy. This process can pinpoint locations ranging from a street level to city or country-wide areas. In recent years, the scientific community has shown a growing interest in using social media geolocation for disaster management research. Various studies have adopted methods using images, text, and combined approaches to improve the accuracy of geolocation identification.

In an image-based approach,(Murgese et al., 2022) proposed a solution to estimate images' locations by implementing Focal Modulation Networks for predicting the geolocation. The authors used geolocated images from Flickr2 and Mapillary3 to evaluate the proposed approach. Mapillary is an open-source map that provides street-level scenes. The focus of the study was the Barcelona City, Spain. Around 70% of Mapillary data, which had street-level imagery similar to Google Street View, was successfully pinpointed at neighborhood level. However, the proposed approach did not use social media data, and the precision of geolocation prediction was limited to the neighborhood level.

A multimodal (combining visual and textual features) approach was introduced by (Tahmasebzadeh et al., 2023). In the work, the author predicted the geolocation of news photos. To implement the approach, the authors also introduced a new dataset called Multimodal Geolocation Estimation of News Photos (MMG-NewsPhoto). The work demonstrated that the multimodal approach outperformed the unimodal.

Related works on text-based geolocation use several approaches, for instance, Named Entity Recognition and Location Indicative Words, to infer the geolocation candidate (Utomo et al., 2018). For example,(Middleton et al., 2018), implemented the Named Entity Recognition algorithm to predict geolocation from social media data. In the experiment, they used Twitter4 and Flickr as the experimental dataset. The works were evaluated qualitatively and quantitatively. However, the experimental result demonstrated low results with a high radius from the location.

Another technique called Location Indicative Words, adopted by (Han et al., 2012), used a dataset of cities in the experiment. They developed and evaluated the approach on the English dataset. The experimental result demonstrated that the feature selection improved information gain and outperformed state-of-the-art geolocation prediction methods by 10.6% in terms of accuracy. Also, the result showed the mean and median of prediction error distance by 45 km and 209 km. Despite the innovative technique, the precision deviation was still high.

1https://publications.jrc.ec.europa.eu/repository/handle/JRC124081
2https://www.flickr.com/
3https://www.mapillary.com
4https://twitter.com

In a recent study,(Scalia et al., 2022) proposed a context-aware approach to geolocate emergency-related social media posts. The system implemented post metadata and post network of relationships. The experimental results showed that around 35% of images geolocation inferences were within a 500-meter radius from the ground-truth location. About 62% of images were predicted within 12 km. However, the results still had a high radius from the actual location.

Text-based geolocation prediction research often includes text from social media posts. The other alternative is by extracting the text from an image and using the extracted text to retrieve the location. (Firmansyah et al., 2023) proposed an automated text-extraction pipeline to predict the geolocation of an image. The work used the text in a social media image as input for a text-based geolocation prediction algorithm. The experiment incorporated 1,752 images curated from four different disasters in the world (Croatia Earthquake, Catania Floods, Central Europe Floods, and Haiti Earthquake). The experimental results demonstrated that the number of original images that could be successfully geolocated within the bounding boxes (affected area) varies between 9% and 11% of the original dataset, depending on the geolocation algorithm used. However, the work was still in the preliminary phases and had a country-level precision.

In a hybrid technique, (Ravi Shankar et al., 2019) introduced a platform called Crowd4EMS. This platform collects, analyzes, and geolocates social media information by combining crowdsourcing and automatic methods. They used data related to the Amatrice Earthquake in 2016 to evaluate the platform, coming from Flickr, Twitter, and Youtube. The work demonstrated that the suggested approach could better support the crowdsourcing communities in providing high-precision geolocation in the context of disaster response. However, the work also shows that crowdsourcing the geolocation of social media images is a complex and time-consuming task, which produces unpredictable delays that could prevent the information from being available in the immediate aftermath of the event.

In addressing the challenge of combining expert analytical capabilities with the scalability of crowdsourced efforts, the study introduces "GroundTruth," an online system designed to augment expert tasks with crowd participation. This system employs three distinct shared representations—a diagram, grid, and heatmap—to facilitate real-time collaboration between experts and crowds in the complex task of image geolocation, a critical process utilized by professional journalists and human rights investigators. Through a mixed-methods evaluation involving 11 experts and 567 crowd workers, the findings demonstrate GroundTruth's effectiveness in aiding image geolocation tasks, while also highlighting the interaction challenges and successful strategies between experts and non-experts. The research contributes valuable insights into the design of shared representations for enhancing visual search and sensemaking tasks, extending its implications beyond image geolocation to broader applications (Venkatagiri et al., 2019).

RESEARCH QUESTIONS

This paper proposes that the timing and accuracy of the overall geolocation process of social media posts using crowdsourcing and automatic methods can be improved by prioritizing those images deemed easy to geolocate. In this way, easy-to-geolocate images are presented earlier to the crowd, preventing delays from difficult or even impossible-to-geolocate social media posts. This process requires the implementation of an image classifier to detect the level of difficulty associated with the process of geolocating an image.

This work answers the following research questions (RQs):

- **Can the proposed classifier automatically predict the difficulty of geolocating an image?** (RQ1)
- **Does favoring images easy-to-classify improve the overall geolocation process?** (RQ2)

DATASET, METHODOLOGY, AND EXPERIMENT

This section provides an overview of the dataset and methods utilized in the experiment, followed by a comprehensive explanation of the experimental design. To answer the proposed research questions, we devised three major tasks: In the first task, images extracted from a reference dataset are presented to a crowd of domain experts that labels each image with a grade representing the difficulty level of geolocalization. To answer the RQ1 stated above, the second task aims at training a deep learning model for classification using the data annotated in the first task. Finally, the third task consists of a domain adaptation experiment where the same crowd of experts is asked to geolocate previously unseen images annotated by the model, thus answering the second research question (RQ2).

Table 1. Summary of the image geolocating results considering just the total number of geolocatable records

Records	Catania floods		EU floods		Haiti earthquake		Croatia earthquake		All datasets	
	N	%	N	%	N	%	N	%	N	%
Geolocatable	113	100	271	100	1,462	100	627	100	2,473	100
Geolocated	70	61.9	214	79	961	65.7	549	87.6	1,794	72.5
Timeout	43	38.1	57	21	501	34.3	78	12.4	679	27.5

Dataset

As a reference dataset for this research, we used the images geolocated by GIS (Geographic Information System) experts for the study EMSV070 - Reference Datasets for SM (Social Media) Image Geolocation in DRM (Disaster Risk Management) requested within the Copernicus Emergency Management Services framework. The dataset contains geolocated images and text collected during four different disasters: (1) Croatia Earthquake, 2020 (30.7%); (2) EU Floods, 2021 (3.9%); (3) Central European floods, 2021 (11.9%); and (4) Haiti Earthquake, 2010 (53.5%) (see Table 1).

Preferably, the damage in the image (flooded point, damaged building, etc.) was selected for geolocation. There are 5,430 records in the four datasets together. Of these, almost 13% (693) were repeated (non-analyzed), and about 87% (4,737) were analyzed. The details of 4,737 records were 2,264 non-geolocatable and 2,473 geolocatable. Out of 2,473 geolocatable records, we opted for 1,794 records, while 679 records were excluded because they timed out during image access.

Within the analyzed records, the non-geolocatable images were identified and discarded by GIS experts. Reasons for being considered as non-geolocatable are:

1. The URL of the image does not work, or the URLs of both the image and tweet do not work.
2. The information visible on the image does not allow its geolocation since it is a mosaic of images or a painting, or the image has bad quality (e.g., it is blurry, dark, etc.), or the image does not have visible, recognizable elements to be geolocated unless external information helps in its geolocation.
3. The image could be geolocated only at the city level.
4. The image does not correspond to any of the disaster events under study.
5. The image corresponds to regions outside of the study areas.

The geolocation was performed by GIS experts considering the following:

1. The information provided by the images themselves.
2. The text contained in the tweets.
3. The images and/or videos contained in the tweets.

Google Street View, Mapillary, Google Earth Pro, and Google Maps were used for localizing the places, recognizing the visible elements in the images, and/or for the subsequent image geolocation. Additionally, the already recognized images during the execution of the task were also considered for the geolocation of new images. Mainly, the geolocation was performed by:

1. Focusing, when possible, on the geographic information specified in the text of the tweets.
2. Selecting in the image some location relevant to the disaster event under study (flooded area or damage due to an earthquake), which was expected to be easily geolocated with the help of Google Street View or Mapillary, etc.
3. Selecting some feature/s in the image that help locate the photo the viewers used.
4. Providing the [Latitude] and [Longitude] for the selected location.

Consequently, in our work, we used images that were geolocated by the experts at least at the level of city and in 90% of cases with a precision of 10 to 100 meters. Figure 1 depicts the geolocation accuracy distribution of four different disasters.

Accuracy (m)	Catania Floods		EU Floods		Haiti Earthquake		Croatia Earthquake		All datasets	
	N	%	N	%	N	%	N	%	N	%
<10	65	92.9	173	80.8	647	67.3	532	96.9	1,417	79
<100	1	1.4	25	11.7	213	22.2	17	3.1	256	14.3
<1000	4	5.7	16	7.5	101	10.5	0	0	121	6.7
>1000	0	0	0	0	0	0	0	0	0	0
Total	70	100	214	100	961	100	549	100	1,794	100

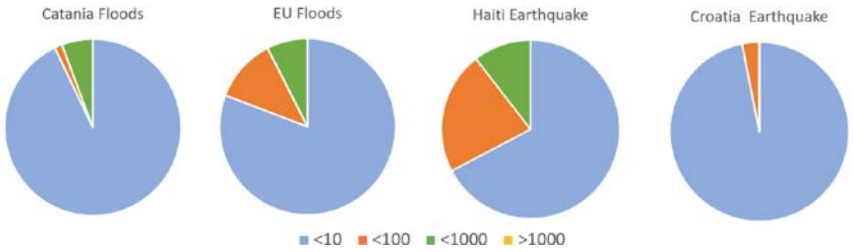


Figure 1. Distribution of the accuracy values(%) assigned to the images geolocated in each of the four datasets

We asked experts in the domain of Disaster Response, specifically the European network VOST Portugal (VOSTPT) to classify the images of the reference dataset, each based on the difficulty of geolocating them. VOSTPT is a community of citizens from different countries accross Europe that has an objective to create a more prepared, more informed, and more resilient society to natural disasters and other events 5. To label the images, four different levels were used: (1) Easy, (2) Medium, (3) Difficult, or (4) Impossible to geolocate (see Figure 2). Each image is labeled at least by five different raters.



Figure 2. The example of geolocation difficulty level for image

The crowdsourcing of the images was carried out via the Pybossa framework (Gonza´lez et al., 2020). Specifically, the execution of this task uses an instance of Pybossa hosted at the Citizen Science Center Zurich, a joint initiative created by the University of Zurich and ETH Zurich6. The application is called the Citizen Science Project Builder (CSPB).

5https://vosteurope.org/
6https://citizenscience.ch/en/

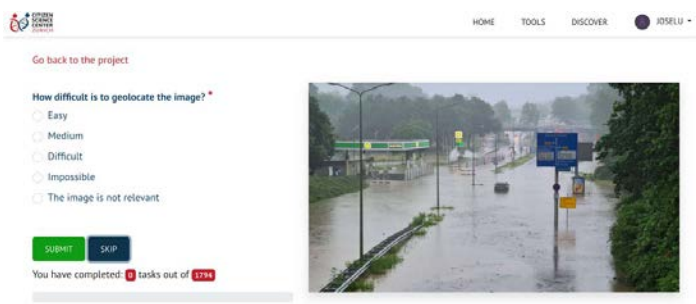


Figure 3. Screenshot of the Citizen Science Project Builder used to crowdsource the labeling of images depending on the level of difficulty to find its location

Figure 3 shows the user interface implemented for crowdsourcing the geolocation difficulty level of each picture in the dataset. The crowd is asked to determine the level of difficulty to geolocate the image on the right. The user interface also enables users to answer two different conditions stating whether it is impossible to geolocate or the image is not relevant. If the users were not sure about their choice, they could also skip the task.

As for the dataset for the second task, i.e., training our deep learning model, Majority Voting and Dawid-Skene (Dawid & Skene, 1979) consensus mechanisms were adopted to aggregate the annotations collected in the first task. We detail both mechanisms in the experimental results section.

Training a CNN to Predict the Level of Difficulty in Geolocating a Given Photo

Our second task is to create a classifier to classify the difficulty level to geolocate a given photo automatically. To this end, we implemented a Convolutional Neural Network (CNN) and the methodology for training the CNN consisted of the following steps:

First, we imported images from the annotated training dataset. We loaded images and distributed them among the classes "Easy" and "Difficult". In the beginning, we decided to use all the labels existing in the annotations, but after some preliminary analysis, we decided to reduce the complexity of the classification. Hence, we removed the Medium samples from the dataset and merged Difficult and Impossible into Difficult to make the task a binary classification between Easy (Easy) and Difficult (Difficult or Impossible). The reason for this choice was to focus on accurately identifying True Positive values (the Easy class).

We then deduplicated the data using hashing algorithms[7] that are particularly good at finding exact duplicates, as well as CNNs, which are also good at finding near duplicates. We then used some standard data analysis to check the distribution of images. Figure 4 shows the distribution of the images after the deduplication. The classes were well-balanced.

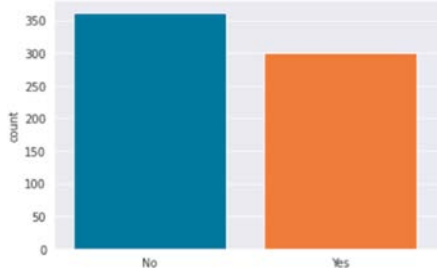


Figure 4. The almost-balanced distribution of images after deduplication

We split the dataset with a training, validation, and test ratio of 70% (448 images), 20% (150), and 10% (61), respectively. The data preparation phase included data augmentation and pre-processing. The data augmentation increased the diversity of our training set by applying random (but realistic) image rotation and image flipping (horizontally). Figure 5 shows an example of data augmentation.

[7]https://pypi.org/project/imagededup/



Figure 5. Image augmentation examples. The augmentation flips and rotates the images at certain degrees.

This step basically multiplies the number of images used in our training. Every image was augmented by being flipped and rotated. After freezing the all layers, the first 25 epochs to train the neural network. The next 25 epochs were used to fine-tune the neural network. The number of epochs was inspired by a similar experiment on leveraging a pre-trained model using social media images for natural disaster (Firmansyah et al., 2022; Ofli et al., 2020) and considering several factors such as dataset size, overfitting prevention, and learning rate.

EfficientNet 8 is among the most accomplished models that reach state-of-the-art accuracy on both ImageNet 9 and common image classification transfer learning tasks. The model was performing better than other models and could be chosen as the best option (Alam et al., 2020). In our experiment, we chose the EfficientNet-B6 due to its good balance between computational demand and model performance and because it uses an image size (528x528 resolution) similar to the ones found on social networks. Aware of the fact that training a deep learning model from scratch with a relatively small dataset might lead to suboptimal results, we implemented transfer learning. It involved starting with a pre-existing model trained on a large dataset and adapting it to our task. This practice reduces the data needed, reduces training time, and improves the model's performance on smaller datasets.

To customize the pre-trained model for our specific task, we performed fine-tuning on the last 66 layers of the EfficientNet-B6 in a two-steps process: first freezing all layers and training only the top layers, and then unfreezing the last 66 layers (out of 666). A relatively large learning rate (1e-2) was used for the first step, while the second step fit the model using a smaller learning rate. Fine-tuning involves adapting the weights of a pre-trained model to new data, allowing us to leverage both the general feature extraction learned during pre-training and the specific patterns found in our disaster imagery.

In conclusion, the experiment sought to leverage the power of EfficientNet-B6 and the robustness of noisy student pre-training by fine-tuning the model to our specific classification. Noisy student is a semi-supervised learning method where a model trained on labeled data (teacher model) produces pseudo-labels for unlabeled data. The pseudo-labels are then combined with noisy data and labeled data to train another model called the student model (Xie et al., 2020). Figure 6 shows a comparison of accuracy while testing different weights. It compared two EfficientNet-B6 models with two different weight (ImageNet and Noisy Student). The EfficientNet-B6 model used a training dataset of 448 images as explained above.

The base classifiers were trained with the objective of decreasing the cross-entropy loss. Cross-entropy measures how well a classifier approximates the probabilities of its predictions. The code generated to create and evaluate the classifier is accessible as open-source10.

We ran two experiments with two datasets originating from Majority Voting (MV) and Dawid-Skene (DS) consensus algorithms and compared the performances of the two models. The model trained with the DS dataset presented better results than the one with Majority Voting, as we expected since DS weighs the raters' answers based on their error rates computed by the same algorithm. That is, the annotations of the raters who perform better (i.e., make fewer errors) have more influence on the consensus calculation. Thus, we expect to have a more reliable consensus with DS. Figure 7 demonstrates the comparison of the training accuracies with MV and DS datasets and their evolution after being fine-tuned (the fine tuning started at the vertical green line).

8https://paperswithcode.com/method/efficientnet
9https://paperswithcode.com/dataset/imagenet
10https://github.com/hafizbudi/Geolocation AI Crowdsourcing

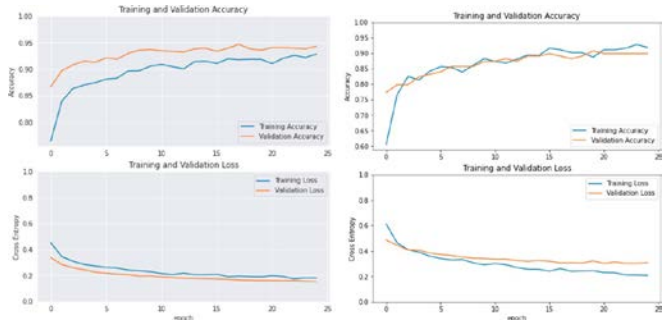


Figure 6. ImageNet (left) and Noisy student (right) weight performance comparison on model training

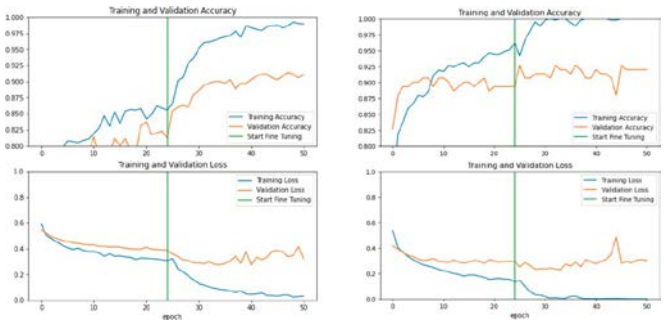


Figure 7. The comparison of model training performances between Majority Voting (left) and Dawid-Skene (right) consensus training datasets.

This preliminary result shows that classifiers could estimate the difficulty of geolocating a social media image. This step is critical to filter out images that are impossible to geolocate and to optimize the ratio of images geolocated per person right after a disaster happens when the time is crucial. Thanks to this methodology, easy images could be scheduled with higher priority than those difficult to geolocate, saving invaluable time for disaster management. We assess the validity of our hypothesis in the following section. Figure 8 shows examples of classifications correctly inferred by the model.

Domain Adaptation Classification Experiment

A central hypothesis in this research is that the photos which are automatically classified as Easy to geolocate should be geolocated with higher precision and faster by humans than the images classified as Difficult. The tool selected for the crowd annotation experiment is Citizen Science Project Builder 11 adopting the user interface design implemented within the E2MC EU project that facilitates manual geolocation (Ravi Shankar et al., 2019).

To answer our RQ2, we set up a domain adaptation classification experiment (Gopalan et al., 2015) where the model described in the previous section classifies unseen photos and human experts geolocate them and provide the precision of the location. Figure 9 and Figure 10 show screenshots from the Citizen Science Project Builder tool with the user interface for geolocating social media images. The interface combines the following: (1) initial estimation of the position of the image extracted from the analysis of the associated text to the photo, which allows participants to start refining the position rather than geolocating an image from scratch; (2) incorporation of Google Map and Google Street View which allow matching the image from social media, with images from Google Street View. When the user locates the photo on Google Street View and clicks the Submit button, the geolocation (longitude and latitude) is submitted automatically within a precision of a few meters. The user interface allows to annotate the precision of the geolocation with three different levels, i.e. the precision the user was able to achieve in the geolocation process.

Figure 11 shows the three levels of precision: i) High precision means the user found the exact location by matching it with the referenced dataset's latitude and longitude. ii) Medium precision means the user knows the image was taken close to a position, such as a street, road, path, or similar, but it could not be matched the referenced one. iii) Low precision means the image is geolocated only at the city level.

11https://lab.citizenscience.ch/en/tools/projectbuilder



Figure 8. Examples correctly classified by the model. The text in the upper of each image corresponds to image geolocation difficulty level

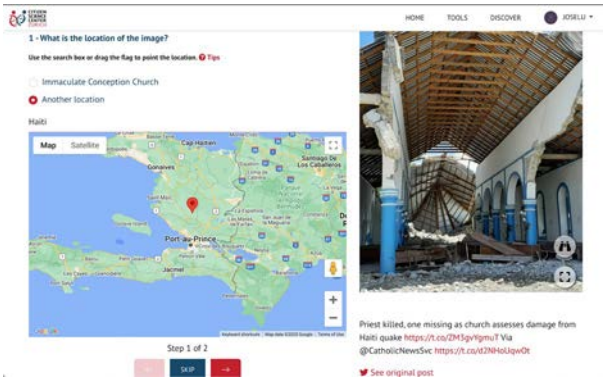


Figure 9. Screenshot of the Citizen Science Project Builder used to crowdsource geolocation of the photos (Photo from Haiti earthquake).



Figure 10. Screenshot geolocation using Google Street View (Photo from Croatia earthquake)

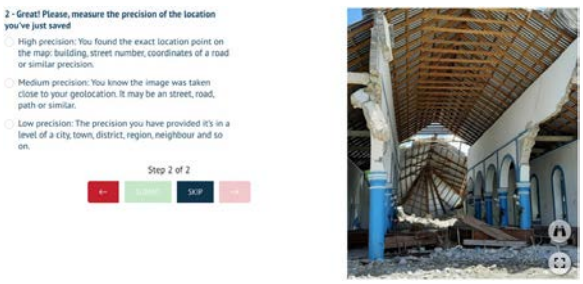


Figure 11. Screenshot CS Project Builder, final step geolocation precision levels.

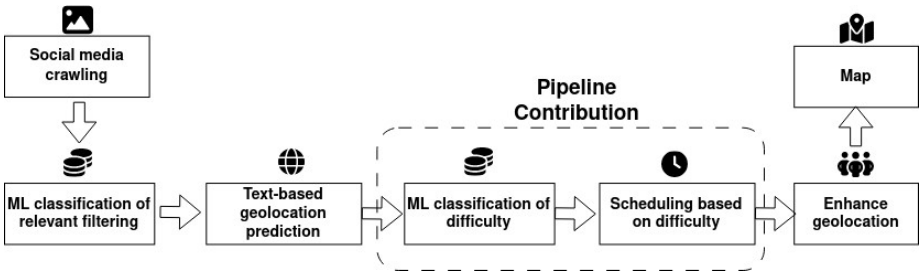


Figure 12. The pipeline of social media images geolocation

Proposed Pipeline

The most common approach to get the geolocation is by extracting location information tagged to the social media post using GPS. Our approach is completely different from that context for two reasons. First, our approach minimized the crowd effort by automatically estimating the geolocation difficulty using machine learning. Second, the approach incorporated crowd to accelerate and improve the image geolocation step.

To make use of the geolocation enhancement for the social media data proposed in this paper, we provide a novel pipeline which extends previous work from Crowd4EMS (Ravi Shankar et al., 2019) and SMDRM (Lorini, Panizio, & Castillo, 2022) by adding two additional processes. Figure 12 shows the overview of a proposed pipeline which is composed of six processes. The dashed line indicates the two additional processes integrated into the pipeline:

- **Social media crawling** fetches informative posts from social media by selecting appropriate keywords which depend on the type of disaster. The initial step is to get the social media post by formulating the appropriate search query (Alam et al., 2018; Lorini, Panizio, & Castillo, 2022).
- **ML classification of relevant filtering process** filters the irrelevant images and keeps the relevant images crawled from the previous process. For example, images containing cartoons, banners, news logos, and edited images are considered irrelevant (Alam et al., 2018).
- **Text-based geolocation prediction** uses social media posts to predict the initial geolocation of images. The main goal of the process is to provide initial geolocation information based on social media posts. The text-based geolocation prediction is used to prevent the crowds to find an image location from scratch, when possible.
- **ML classification of difficulty** predicts the difficulty level to geolocate an image into predetermined categories (i.e. easy, medium, difficult, etc). The process is critical because the institution's resources could put attention to the easier-to-geolocate and more relevant images first instead of to all images. For example, images that are impossible to geolocate should not be sent to the crowd. The process of classification difficulty is done by implementing automated machine learning (ML) described previously in Section Training a CNN to predict level of difficulty in geolocating a photo.
- **Scheduling based on difficulty** chooses images in a desired order of difficulty level to be sent to the crowd. In certain situations like disasters, the time constraint may be more strict. When time is considered an essential dimension for geolocating an image, the first responders organization should prioritize images that need to be sent to the crowd earlier (e.g., Easy before Difficult images as labeled by the automated machine learning).
- **Enhance geolocation** incorporates crowd to improve the geolocation precision.
- **Map** visualizes the geolocation enhancement on a map.

EXPERIMENTAL RESULTS

This section provides the experimental results from the three different tasks presented in Section Dataset, methodology and experiment. The aim of this section is to answer the two research questions.

As the pilot study, we first worked with 1,794 photos annotated by the VOSTPT community using the interface presented in Figure 3, with the redundancy of at least five annotations per image. In total, 8,970 classifications were performed by the VOSTPT community were each image was classified by the level of geolocation difficulty, i.e. Easy, Medium, Difficult and Impossible. With the annotation dataset in hand, we initially evaluated the inter-rater

agreement. Fleiss' Kappa12 is a common statistical measure to assess the reliability of agreement between raters in the context of categorical ratings, like our case. We computed the Fleiss' Kappa for the VOSTPT dataset using the Python library Crowdnalysis (Cerquides & Mu¨la^yim, 2022). The value was 0.288 which indicates a Fair agreement on the kappa scale among the VOSTPT community for this dataset. The low value of fleiss kappa suggests that the task of assigning a geolocation difficulty level to a social media image is a complex task and the annotation result might be noisy. Also, there was a probability that the annotator would answer the level of difficulty randomly. Considering that issue in the experiment, we decided to choose Dawid-Skene (DS) model instead of majority voting (MV) since DS also models individual annotator behavior. We further explain this choice below.

To generate the dataset for the next step, i.e., the training of the CNN, we needed to find a consensus between the five annotators for each image. The Crowdnalysis library, besides the standard Majority Voting, provides advanced probabilistic methods of consensus (Cerquides et al., 2021) such as the seminal Dawid-Skene model (Dawid & Skene, 1979) (DS). The DS method allows modelling individual annotator behaviour, thus with enough data, it yields more reliable consensus results which can prove crucial in disaster management scenarios, as mentioned earlier.

Table 2. The percentage of warnings in consensuses with Dawid-Skene (DS) and Majority Voting (MV) models.

	Easy	Medium	Difficult	Impossible	Not Relevant
DS	**4.4%**	**5.2%**	**9.7%**	**0.7%**	0.43%
MV	57.2%	58.8%	10%	2.7%	**0%**

Table 2 shows the percentage of warnings for each consensus result. The warning was a condition where there was only a < 0.1 difference between the probabilities of the top and second-best estimated consensus classes for the photo. The boldface values indicate the lower values between models trained with DS and MV consensus datasets.

Table 3 shows the confusion matrix result of evaluating the model trained with the DS dataset. The model exhibits 0.87 accuracy, 0.88 precision, and 0.88 recall. Assessing the difficulty of geolocating an image is challenging even for humans, as demonstrated by the Kappa value above. However, the evaluation of the proposed model shows that it is possible to automatically predict the difficulty of geolocating an image. This result answered the first Research Question (RQ1).

Table 3. Confusion Matrix of the CNN model with binary classification

	Predicted	
Actual	Difficult	Easy
Difficult	29	4
Easy	4	24

After demonstrating the feasibility of predicting the difficulty of geolocating an image from social media, the next question to address is whether images classified as easy contribute to the enhancement of the overall geolocation process. The main focus is to investigate whether images labeled as easy result in increased precision and reduced time for geolocation during the crowdsourcing process. Figure 13 shows that most of the images predicted as easy are gelocated with high precision (horizontal axis = 2) and within 300 seconds. On the other hand, Figure 14 shows that images predicted as difficult are geolocated with low precision (horizontal axis = 0). Overall, prioritizing easy-to-geolocate images accelerates the geolocation time and enhances precision. The detail of the results is explained below.

We then asked experts to geolocate 39 images. Those images were classified using previously trained model. From that automated classification,about 22 images were predicted as Easy. Figure 13 shows that most of the images classified as Easy are geolocated with high precision (within a few meters). Most of the images geolocated with high precision were geolocated in less than 4 minutes. If we assume an average of 3 minutes per photo, a person could potentially geolocate 20 images per hour. Consequently, in a group of 10 volunteers, we could geolocate 200 images per hour with high precision. On the other hand, 17 images were predicted as Difficult. Figure 14 shows the precision and time spent on geolocating images classified as Difficult to geolocate. We could see how most of them are geolocated with low precision within the city level.

Although this is a preliminary result based on a limited number of images, it clearly shows that prioritizing images easy-to-classify during the crowdsourcing process could potentially augment the overall geolocation process, by increasing the precision and reducing the time needed to geolocate the images.

12https://en.wikipedia.org/wiki/Fleiss_kappa

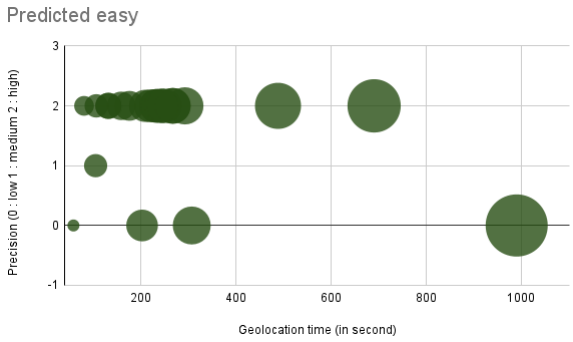


Figure 13. Geolocation level of images predicted as Easy to geolocate. The horizontal axis and the size of the radii represent the average time spent by users to geolocate the photo

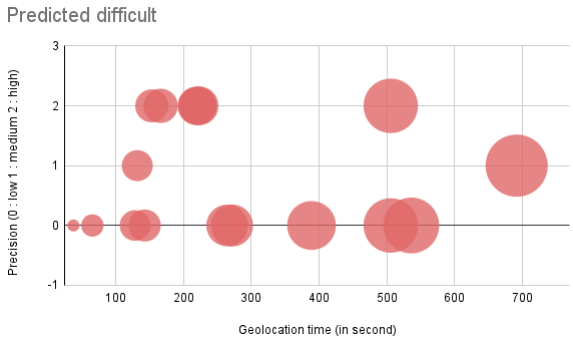


Figure 14. Geolocation level of images predicted as Difficult to geolocalize. The horizontal axis and the size of the radii represent the average time spent by users to geolocate the photo

CONCLUSIONS AND FUTURE WORK

Finding the location of social media images related to a disaster is a key process to make the data actionable, i.e., data that can be used to enable better-informed decisions. Geolocation of social media images is a task usually performed by humanitarian communities such as Stand By Task Force, GISCorp, or VOSTPT. It is a well-known fact that crowdsourcing the geolocation of social media images is always a difficult task which slows down the process of assessing the impact of a natural or man-made disaster in the immediate aftermath of the event. This preliminary research activity demonstrated that it is indeed possible to automatically predict the difficulty of geolocating an image, thus allowing to filter out impossible-to-geolocate images and giving priority to images that can be geolocated within a few minutes with high precision. Consequently, the proposed approach would enable a more efficient and timely analysis of social media data which is critical in a disaster.

Finally, regarding future work, first, the proposed approach needs to be evaluated in a real-time crisis response to demonstrate the potential of using social media data to assess the impact of disaster within 24 hours. Second, we will increase the number of geolocated images to answer the research questions in a robust manner.

ACKNOWLEDGEMENTS

This work was funded by the Indonesia Endowment Fund for Education (LPDP) and partially supported by the European Union's Horizon research and innovation programme Crowd4SDG and K-HealthInAIR projects (under grant agreements ID 872944 and 101057693, respectively) and by the grant PID2022-136787NB-I00 funded by MCIN/AEI/10.13039/501100011033.

REFERENCES

Alam, F., Ofli, F., & Imran, M. (2018). Crisismmd: Multimodal twitter datasets from natural disasters. *Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM)*.

Alam, F., Ofli, F., Imran, M., Alam, T., & Qazi, U. (2020). Deep learning benchmarks and datasets for social media image classification for disaster response. *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 151–158. https://doi.org/10.1109/ASONAM49781.2020.9381294

Cerquides, J., & Mu¨laˆyim, M. O. (2022, January). Crowdnalysis: A software library to help analyze crowdsourcing results. https://doi.org/10.5281/zenodo.5898579

Cerquides, J., Mu¨laˆyim, M. O., Herna´ndez-Gonza´lez, J., Ravi Shankar, A., & Fernandez-Marquez, J. L. (2021). A conceptual probabilistic framework for annotation aggregation of citizen science data. *Mathematics*, *9*(8). https://doi.org/10.3390/math9080875

Dawid, A. P., & Skene, A. M. (1979). Maximum Likelihood Estimation of Observer Error-Rates Using the EM Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *28*(1), 20–28. https://doi.org/10.2307/2346806

Firmansyah, H. B., Cerquides, J., & Fernandez-Marquez, J. L. (2022). Ensemble learning for the classification of social media data in disaster response. *ISCRAM 2022 Conference Proceedings – 19th International Conference on Information Systems for Crisis Response and Management*.

Firmansyah, H. B., Fernandez-Marquez, J. L., Cerquides, J., Lorini, V., Bono, C. A., & Pernici, B. (2023). Enhancing disaster response with automated text information extraction from social media images. *The 9th IEEE International Conference on Big Data Computing Service and Machine Learning Applications*.

Gonza´lez, D. L., alejandrodob, therealmarv, Keegan, M., Mendes, A., Pollock, R., Babu, N., Fiordalisi, F., Oliveira, N. A., Andersson, K., & et al. (2020). Scifabric/pybossa: V3.1.3. https://doi.org/10.5281/zenodo.3882334

Gopalan, R., Li, R., Patel, V. M., & Chellappa, R. (2015). https://doi.org/10.1561/0600000057

Han, B., Cook, P., & Baldwin, T. (2012). Geolocation prediction in social media data by finding location indicative words. *Proceedings of COLING 2012*, 1045–1062.

Huang, B., & Carley, K. M. (2020). A large-scale empirical study of geotagging behavior on twitter. *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 365–373. https://doi.org/10.1145/3341161.3342870

Lorini, V., Panizio, E., & Castillo, C. (2022). Smdrm: A platform to analyze social media for disaster risk management in near real time. *In Workshop Proceedings of the 16th International AAAI Conference on Web and Social Media*.

Lorini, V., Rufolo, P., & Castillo, C. (2022). Venice was flooding ... one tweet at a time. *Proc. ACM Hum.-Comput. Interact.*, *6*(CSCW2). https://doi.org/10.1145/3555107

Middleton, S. E., Kordopatis-Zilos, G., Papadopoulos, S., & Kompatsiaris, Y. (2018). Location extraction from social media: Geoparsing, location disambiguation, and geotagging. *ACM Trans. Inf. Syst.*, *36*(4). https://doi.org/10.1145/3202662

Murgese, F., Costa-Alcaina, G., Mu¨laˆyim, M. O., Cerquides, J., & Fernandez-Marquez, J. L. (2022). Automatic outdoor image geolocation with focal modulation networks. *Artificial Intelligence Research and Development: Proceedings of the 24th International Conference of the Catalan Association for Artificial Intelligence*.

Ofli, F., Alam, F., & Imran, M. (2020). Analysis of social media data using multimodal deep learning for disaster response. *arXiv preprint arXiv:2004.11838*.

Ravi Shankar, A., Fernandez-Marquez, J. L., Pernici, B., Scalia, G., Mondardini, M. R., & Serugendo, G. (2019). Crowd4EMS: A crowdsourcing platform for gathering and geolocating social media content in disaster response. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-3/W8*, 331–340. https://doi.org/10.5194/isprs-archives-xlii-3-w8-331-2019

Rufolo, P., Muraro, D., & Lorini, V. (2021). *Social media image analysis in the immediate aftermath of the 2020 beirut blast*.

Scalia, G., Francalanci, C., & Pernici, B. (2022). CIME: Context-aware geolocation of emergency-related posts. *GeoInformatica*, *26*(1), 125–157. https://doi.org/10.1007/s10707-021-00446-x

Tahmasebzadeh, G., Hakimov, S., Ewerth, R., & Mu¨ller-Budack, E. (2023). Multimodal geolocation estimation of news photos. In J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, & A. Caputo (Eds.), *Advances in information retrieval* (pp. 204–220). Springer Nature Switzerland.

Utomo, M. N., Adji, T. B., & Ardiyanto, I. (2018). Geolocation prediction in social media data using text analysis: A review. *2018 International Conference on Information and Communications Technology (ICOIACT)*. https://doi.org/10.1109/icoiact.2018.8350674

Venkatagiri, S., Thebault-Spieker, J., Kohler, R., Purviance, J., Mansur, R. S., & Luther, K. (2019). Groundtruth: Augmenting expert image geolocation with crowdsourcing and shared representations. *Proc. ACM Hum.-Comput. Interact.*, *3*(CSCW). https://doi.org/10.1145/3359209

Xie, Q., Luong, M.-T., Hovy, E., & Le, Q. V. (2020). Self-training with noisy student improves imagenet classification. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10684–10695. https://doi.org/10.1109/CVPR42600.2020.01070

Social Media Analysis in Sudden Onset Disasters and its Usefulness for Decision Makers - Excerpt of a Scoping Review

Francesca Müller*
University Wuppertal
framueller@uni-wuppertal.de

Sylvia Bach
University Wuppertal
sbach@uni-wuppertal.de

Frank Fiedrich
University Wuppertal
fiedrich@uni-wuppertal.de

ABSTRACT

Problem: Despite the potential of artificial intelligence for rapid social media analysis, practical implementation remains limited. **Approach:** We conducted a scoping review using iteratively developed search terms in two electronic databases. We then conducted a descriptive and thematic analysis to capture the key concepts, findings and methods in the literature. **Results:** After reviewing 130 papers, 27 papers were identified that met our inclusion criteria. Four key themes emerged, including tool development, tool performance comparison, gaining insights from social media data, and usefulness for decision makers. **Conclusions:** While acknowledging the importance of integrating social media into crisis and disaster management, there is a lack of empirical evidence on data reliability and preparation for decision makers. More empirical studies on social media analytics, specifically focusing on usefulness for crisis and disaster management, are recommended.

Keywords

scoping review, social media, crisis and disaster management, sudden onset disaster, evidence

INTRODUCTION

Effective disaster relief and management requires a comprehensive understanding of the disaster situation (Fan et al., 2020). As the frequency and intensity of crises and disasters increases, decision makers' knowledge of the situation is becoming ever more important to enhance the effectiveness of response and recovery efforts. Decision makers in the context of crises and disasters are individuals or entities responsible for making critical decisions related to emergency response, recovery, and management. In the digital age, social media (SM) has evolved beyond its role as a platform for personal exchange to become a valuable source of information that can reflect the collective intelligence of the public (Lang et al., 2020). Gaining insights into people's perceptions, reactions and behavior on SM can contribute significantly to crisis and disaster management (Z. Wang & Ye, 2017). Acknowledging the collective intelligence of the public as reflected in social media is a key aspect. This connects with broader discussions on Open Source Intelligence techniques, citizen engagement, and the potential for harnessing distributed knowledge during crisis situations (Pastor-Galindo et al., 2020). However, the sheer volume and variety of information on SM makes manual processing in real time almost impossible (Avvenuti et al., 2015). To overcome this challenge, various artificial intelligence techniques have emerged that offer the possibility of efficiently classifying, analysing and visualizing SM data. This aligns with broader research on the application of AI in disaster response, including natural language processing, machine learning, and data visualization (Abid et al.,

*corresponding author

2021; Fan et al., 2020). The passive collection of information, i.e. without active participation in the conversation or demand for information, allows topics and problems relevant to the population (Ghermandi & Sinclair, 2019). Passive information gathering aims to use existing data sources without disrupting user activity. While numerous studies have examined the technical implementation of these techniques, the integration of SM insights into the overall picture of the situation remains limited in practice due to the complexity of existing algorithms and the lack of comprehensive software solutions. Due to the high number of scientific publications in the field of SM analysis in crises and disasters, there is an increasing need for overviews. However, existing reviews focus unidimensionally on thematic or methodological orientations of the evidence, so that the temporal dimension and usefulness for decision makers are not explicitly considered (Acikara et al., 2023; Luna & Pennock, 2018). This review aims to fill this gap by capturing the available evidence, including the temporal dimension, and providing an overview of the use of passive information gathering from SM for crisis and disaster management in sudden onset disasters. The following research questions guide this review paper:

RQ1.1 How does the number of publications about social media analyses related to the support of decision makers during sudden onset disasters vary over time?

RQ1.2 On what data are the publications on social media analytics related to supporting decision makers during sudden-onset disasters based, taking into account the evolution over time?

RQ2 What methods are used to obtain information from social media to enhance the situation picture for decision makers in sudden onset disasters?

METHODS

A scoping review was used to capture the key concepts, main sources and types of evidence available in the literature on the passive extraction (observing without active communication) of information from SM in sudden onset disasters for decision makers in crisis management teams and consequently to complement the situational picture. Sudden-onset disasters are events that occur suddenly and unexpectedly, usually with little or no warning, and have an immediate and severe impact on communities and infrastructure, i. e. earthquakes and floods. The focus is intended to provide an overview of direct practical implications and options for disaster management decision-makers in time-critical contexts. The review was based on the methodological guidance for scoping reviews according to the Joanna Briggs Institute (JBI) (von Elm et al., 2019) and conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR guidelines) (Moher, 2009). The overarching research project *#sosmap* focuses on the state of the art regarding the passive derivation of information from SM on the psychosocial needs and resources of the population in any crisis and disaster situation. This work-in-progress paper presents a subset of the results in relation to the knowledge gained from SM for decision makers in sudden onset disaster. As scoping reviews aim to provide an overview of the available evidence, regardless of methodological quality or risk of bias, no critical appraisal was undertaken.

Search Strategy

In several iterations, we have developed a search query that covers a range of terms defined by three key topics: (a) SM, (b) decision making and (c) crises and disasters. Two databases were searched from the beginning: IEEE Xplore (until 20.07.2023) and Web of Science (WoS) (until 09.08.2023). When selecting the databases, WoS was chosen because of its in-depth coverage (1945 to the present). IEEE was included due to its technical focus. The review focuses exclusively on primary literature, including published research results in specialized journals, databases, and conference contributions. There is no time limit set, and only English-language literature is considered as no german-language literature was available. The focus on sudden onset disasters was realized by manual title and abstract screening. Table 1 lists the search terms using the syntax for WoS, consisting of the linked topic complexes: extraction of information from SM for decision making in crises and disasters.

Table 1. Key Search Terms

Search terms
(”social media” OR ”user-generated data”) (Title) AND (analy* OR monitor* OR ”social listening”) (Topic) AND (crisis OR crises OR disaster* OR emergenc* OR catastrophe* OR ”extreme event*”) (Topic) AND (”situation* assessment” OR ”situation* aware-ness” OR ”decision mak*”) (Topic) NOT (covid-19 OR marketing) (All Fields)

Study Selection

Scientific publications addressing the three key topics were reviewed, with a focus on passive data collection and

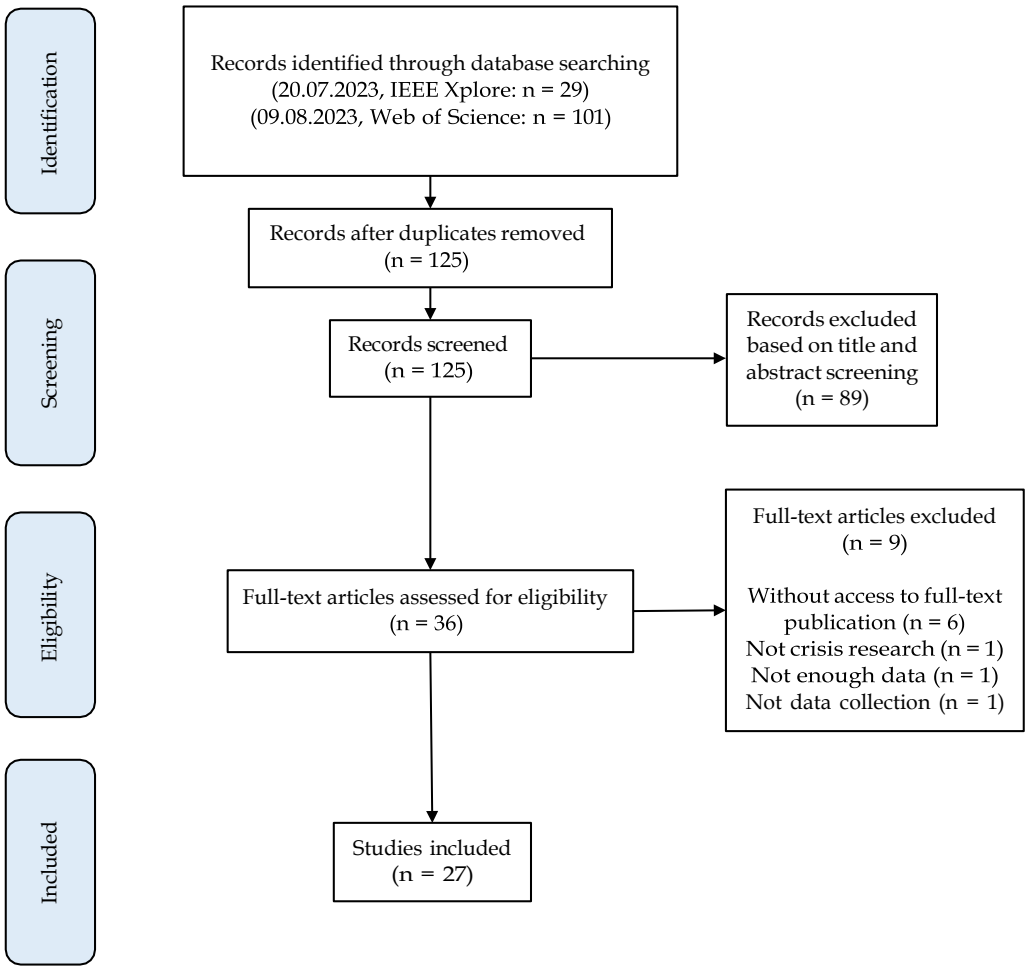


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) Flowchart

analysis via manual, automated, or combined methods. The data sets had to be directly related to sudden onset disasters and support crisis management or complement the situation assessment. Initially, 125 articles were identified, with duplicates removed. Many articles were deemed irrelevant during title and abstract screening, particularly those focusing on active social media engagement or crisis communication. The manual selection of the articles comprised two stages. First, a title and abstract screening was performed to select papers that addressed all three key topics. After the first screening, 36 articles were read in full text. In a second round of full-text screening, data sets were excluded if there was no full-text access and their thematic focus was not sufficiently on all three key topics. A total of 27 studies were identified that were relevant to the research topic. To minimize subjective influences, a second person carried out the two steps on 10 % of the total number of articles. The inclusion and exclusion decisions were identical in 100 % of cases. Figure 1 illustrates the article selection process. The full search protocol is available on request from the authors.

Data Extraction

Details of the included studies can be found in Table 2. Here, for each reference included, the disaster analyzed or the data set analyzed, the social media platforms included for data collection, the data analysis methods used and a concise summary of the findings are compiled. Our in-depth analysis of the included literature included an inductive thematic analysis according to Kuckartz to identify, analyze and present recurring focus areas in the articles (Kuckartz, 2018). The derived focus areas are:

A Development/presentation of tool/framework,

B Performance comparison of tools,

C Gaining knowledge from SM (data insights) and

D usefulness for decision makers.

The presentation of the descriptive summaries takes into account the time dimension.

RESULTS

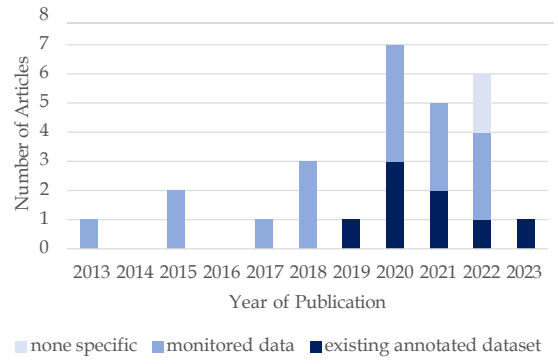
As part of the inclusion and exclusion process, only articles with a publication date after 2013 were integrated. The importance of integrating SM into crisis and disaster management is clearly emphasized by all papers reviewed and beyond. However, very few papers focus on what decision makers really need and adapt the analysis accordingly. Many assumptions are made about what information and presentation methods would be helpful.

Data-specific Consideration

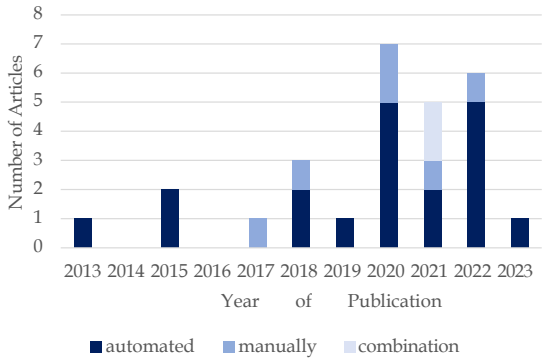
The findings reveal that 63 % of studies relied on monitored data, with annotated datasets comprising 30 %. Notably, the prevalence of existing datasets emerged since 2019. Key datasets included Crisis Benchmark Dataset, CrisisNLP, CrisisLex, MediaEval, European Flood 2013 dataset, Harvard Dataverse, and Hurricane Harvey Twitter dataset, with CrisisNLP being the most used. There's a shift towards utilizing existing and annotated datasets, suggesting a common training basis for models. Figure 2a depicts the trend in data source usage over time. The differentiation was made between already publicly available, annotated datasets of social media data ("existing annotated datasets") and data collected by the authors themselves ("monitored data"). Regarding data formats, 48 % of integrated studies (n_{total} = 27 articles) exclusively used text-based data, while 22 % focused solely on image-based data. Since 2020, there's been a rise in combining text and image data (30 %). The analysis of disaster types showed that 70 % of the studies (n = 19 articles) focused on natural disasters, of which 31.6 % (n = 6 articles) considered several types of natural disasters. Only one study focused on man-made disasters, while 11 % did not focus on a specific crisis and 15 % analyzed several types of disasters. Twitter was the most studied platform (63 %), with Sina Weibo in 7 % of studies. Additionally, 30 % of the studies examined various sources of social media data, including but not limited to Facebook, Instagram, and YouTube. Over time, it can be seen that the consideration of several platforms in one study shows an increasing trend and the singular use of data from Twitter shows a decreasing trend. Since 2020, there's been a decline in studies using Twitter as a data source (-1.6 studies per year).

Method- and Content-specific Consideration

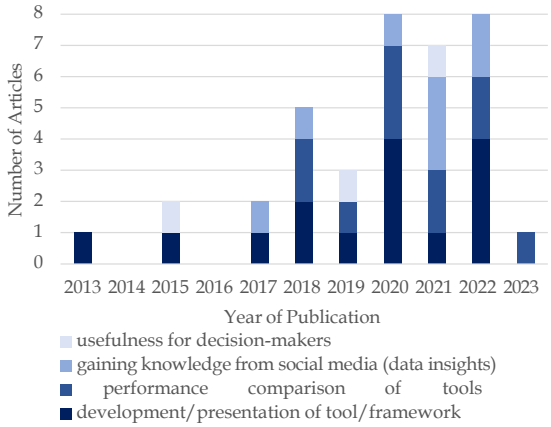
Our analysis revealed that approximately 59 % of articles were case studies. Regarding methodology, 70 % used automated analysis, 22 % manual classification, and only 7 % combined both (Figure 2b). Notable integration of manual and automated analysis is exemplified by Virtual Operations Support Teams (VOST) as highlighted



(a) Data-specific Consideration (Single Assignment)



(b) Method-specific Consideration (Single Assignment)



(c) Content-specific Consideration (Multiple Assignment)

Figure 2. Stacked Bar Charts of the Considered Articles Integrated into the Review

by Fathi and Fiedrich 2022. Article foci included tool/framework development (15 articles), tool performance comparison (11 articles), insights from SM (9 articles), and usefulness for decision makers (5 articles). Figure 2c illustrates the distribution of article foci over time. Named Entity Recognition (NER) was the most commonly used algorithm (6 articles), followed by Convolutional Neural Networks (CNN) and Visual Geometry Group (VGG) architectures (5 articles each). However, the temporal distribution of algorithm usage did not provide further insight into preferences over time.

DISCUSSION OF THE RESULTS

Existing reviews lack a temporal dimension linked specifically to SM in sudden-onset disasters. Therefore, we addressed these gaps by focusing on publications about SM analyses supporting decision makers during sudden onset disasters over time.

RQ1 The analysis reveals a substantial number of studies (63 %) focusing on passive information extraction from SM, primarily from Twitter, with a decreasing trend observed, especially since 2020, possibly due to API changes. Due to the complete switch from Twitter to X and clear restrictions in the API terms and conditions, there could be a further decline and an increased focus on existing data sets in the future. While text-based data remains prevalent, there is a growing trend, particularly since 2020, towards combined analyses of various data formats. Surprisingly, only one study integrated man-made disasters, with a predominant emphasis on natural disasters. Since 2019, there has been an increasing reliance on existing datasets, facilitating more comparative analyses across different crises.

RQ2 The surge in dataset usage is accompanied by a heightened emphasis on comparing different artificial intelligence (AI) techniques, particularly since 2018. However, there is limited consideration of manual classification in comparative or combined methodologies. While concepts like VOST appear promising (Fathi & Fiedrich, 2022), their validation involving decision makers is lacking. Instead, the focus tends to be on understanding crisis situations rather than collaborating with decision makers.

However, our review has limitations. Snowballing was not employed to identify additional literature, potentially restricting the scope. Additionally, screening was not conducted by both examiners for all titles, abstracts, and full texts, possibly leading to omissions. Furthermore, qualitative content analysis for data extraction was performed by both examiners only about 40 % of the time, introducing potential bias. Lastly, our focus solely on sudden-onset disasters may limit the generalizability of our findings.

However, our findings reveal similarities with existing reviews and studies. For example, Karimiziarani (2023) and Kondraganti (2021) emphasize the potential of integrating multimodal data sources, including SM, to improve disaster response efforts. They highlight the importance of ethical considerations and privacy concerns in SM analytics research, advocating for further research to ensure responsible SM data use.

CONCLUSION AND FUTURE OUTLOOK

In this scoping review, we aim to provide a overview of the current state of research regarding passive information extraction from social media (SM) in sudden onset disasters for decision makers. By considering the temporal dimension, we seek to identify the current trends in scientific inquiry.

Looking ahead, it is imperative to shift towards addressing the needs of decision makers during crises and disasters, facilitating the translation of research into practical applications. While numerous articles detail the technical implementation of information extraction from SM, there is a dearth of evaluations targeting the intended user group. Future efforts should prioritize comparative presentations of algorithms and evaluation indicators, such as accuracy and performance, with a clear delineation of their intended use.

This work-in-progress paper offers a glimpse into the ongoing evidence presentation within the research project #sosmap. The forthcoming analysis will delve deeper into an expanded database, implementing an enhanced search strategy focused on deriving psychosocial needs/resources, sentiment, and other factors in various crises.

Table 2. Overview of Selected Article Characteristics

No./ Focus	*Reference*	*Disaster/ dataset*	*Platforms*	*Data analysis methods*	*Summarized results*
1/ B	Alam et al., 2023	Crisis Benchmark Dataset	Twitter	automated (ResNet18, ResNet50, ResNet101, AlexNet, VGG16, DenseNet, SqueezeNet, InceptionNet, MobileNet, EfficientNet)	Optimization of the algorithms is possible through training using consolidated data sets. Based on performance and computational effort, EfficientNet is the best option for real-time applications.
2/ B	Asif et al., 2021	CrisisMMD (CrisisNLP)	multiple	automated (YOLOv4, YOLOv3, VGG16)	With an accuracy of 96 %, YOLOv4 shows better results than YOLOv3 for categorizing images.
3/ A	Avvenuti et al., 2015	Earthquake, Italy, 2012; Flood, Italy, 2014	Twitter	automated (WEKA, NLP, MLR)	Strong earthquakes can be identified by monitoring SM. The algorithm used showed an accuracy of 83.5 % for English tweets and 90.1 % for Italian tweets. In data processing, term clouds contribute to improved situational awareness.

Continued on next page

Table 2 – continued from previous page

No./ Focus	Reference	Disaster/ dataset	Platforms	Data analysis methods	Summarized results
4/ B, C	Dong et al., 2021	Hurricane Sandy, 2012; Hurricane Matthew, 2016; Hurricane Harvey, 2017; Hurricane Michael 2018; Hurricane Dorian, 2019	Twitter	combination of automated + manually (manual classification, LR, NB, DT, SVM, KNN, RF, Adaboost, MNN)	SM is seen as an important platform for gathering information about disasters. It was also found that there is a difference between people's basic needs in natural disasters. The performance of the model adapted to the optimal settings is significantly better than the existing and lexicon-based machine learning models. Moreover, Na¨ıve Bayes is the best performing model in terms of classification accuracy, confusion matrix and ROC-AUC graph in our study.
5/ A, B	Fan et al., 2020	Hurricane Harvey, United States, 2017	Twitter	automated (NER, BERT classifier)	The integration of NER and Google Map Geocoding API enables advanced georeferencing. The fine-tuned BERT-based classifier performs best in classifying tweets into different humanitarian categories. Credible identification of situational information can be implemented through a graph-based clustering approach. In addition, the suggested pipeline can be used to map the development and geographical distribution of disaster events.
6/ C, D	Fathi and Fiedrich, 2022	Floods, Germany, 2021	multiple	combination of automated + manually (ScatterBlogs + VOST)	Information collected by VOST increases situational awareness. The resulting actionable information can contribute to the crisis team's decision-making. Contributions with a potential impact on the health and safety of those affected and image contributions are prioritized higher than others.
7/ C, D	Fromm et al., 2021	Storm Dragi, Germany, 2019; Storm Eberhard, Germany, 2019	Twitter	manually (categorization)	Respondents rate the usefulness of traffic information and eyewitness accounts from private individuals the highest. However, the indication of location and an indicator of relevance and trustworthiness is important. Adaptable filter options are proposed in order to do justice to the disparity between individual assessments of different types of tweets.

Continued on next page

Table 2 – continued from previous page

No./ Focus	Reference	Disaster/ dataset	Platforms	Data analysis methods	Summarized results
8/ A, B	Jixian et al., 2022	none specific	multiple	automated (BiLSTM-Attention-CNN-XGBOOST ensemble model for NN and others)	The BiLSTM-Attention-CNN-XGBOOST model proposed in this paper achieves better accuracy than NB MODEL, Ngrams-TFIDF model, charCNN, wordCNN, LSTM model and RCNN model, regardless of whether it is Chinese or English data from WeiBo, Twitter, Wechat or Facebook platforms.
9/ C, D	Kankanamge et al., 2020	Floods, Australia, 2010-2011	Twitter	automated (capture-understand-present framework, DT, WEKA, IDW)	The use of Twitter is a promising approach, not specifically to gather new situational information, but rather to capture citizens' knowledge. Tweets could be used to identify variations in disaster severity over time. The spatial analysis of tweets confirms the applicability of geographically located messages to delineate severely affected disaster areas.
10/ A, B, D	Lazreg et al., 2019	CrisisLex, CrisisNLP	Twitter	automated (NER, NN, CNTM and others)	The proposed framework for obtaining information from SM provides exactly the information required in 70.09 % of cases.
11/ A, B, C	Z. Li et al., 2017	Floods, United States, 2015	multiple	combination of automated + manually (text matching + manual checking and selection + ArcPy, ArcGIS geoprocessing tools)	There was an 83.4 % match between the map developed based on SM data and official flood maps from the USGS (United States Geological Survey). People share more on SM when they are closer to the event and when the extent of flooding increases during the flood event. Twitter activity reflects the extent of flooding in the relevant geographic area with no obvious time lag.
12/ A, B	X. Li et al., 2018	google image data; Typhoon Ruby/H-agupit, Philippines, 2014; Earthquake, Nepal, 2015; Earthquake, Ecuador, 2016; Hurricane Matthew, 2016	multiple	automated (VGG19, CNN, DDM, GCAM)	The developed Damage Assessment Value in combination with the Damage Detection Map can help assess the severity of damage and identify and prioritize useful information. In addition, the developed tool provides visual explanations for the suggestions made.

Continued on next page

Table 2 – continued from previous page

No./ Focus	Reference	Disaster/ dataset	Platforms	Data analysis methods	Summarized results
13/ B	Lopez-Fuentes et al., 2020	MediaEval 2018	Twitter	automated (double-ended NN and others)	By reducing the number of composite models to 30, the number of networks can be reduced. This makes the solution faster and easier, while at the same time improving the results.
14/ A, B	Mao et al., 2018	power outage (keywords), United States, 2016	Twitter	automated (BiLSTM model coupled with a CRF top layer, and others)	The developed algorithm shows a classification accuracy of 86 % in identifying genuine blackout tweets. In addition, about 20 times more tweets can be georeferenced than are already geotagged. Capturing the number of geo-referenced SM data related to the power outage allows to derive the number of people without power.
15/ A, B	Ning et al., 2020	Floods, United States, 2017; Floods and Florence hurricane, United States, 2018	Twitter	automated (VGG16 with integrated YOLOv3 object detector and a face recognition model)	In general, the developed RIASM system has the potential to be integrated into conventional flood mapping systems to close knowledge gaps and provide additional evidence of the extent of a disaster (total accuracy of flooding photo detection was 93 %).
16/ B	Pereira et al., 2020	European Flood 2013 dataset, MediaEval 2017 (YFCC100m), MediaEval 2018	multiple	automated (DenseNet, EfficentNet, AGDenseNet)	All three models compared showed very similar results. It was found that the number of images available for model training influences the quality of the results. When assessing the water depth, the best results were achieved with the DenseNet model.
17/ A, B	Pohl et al., 2013	Flood, United States, 2011; Bombing, Norway, 2011; Riots, United Kingdom, 2011; Hurricane Irene, United States, 2011	multiple	automated (SOM, AC, 2PG, and 2PGT)	Social multimedia in combination with clustering can be used as part of crisis and disaster management to identify partial events. This allows SM to be integrated into crisis and disaster management without the need for time-consuming manual monitoring. AC and SOM can be used in particular for a quick overview, 2PG and 2PGT for a more user-friendly visualization.

Continued on next page

Table 2 – continued from previous page

No./ Focus	Reference	Disaster/ dataset	Platforms	Data analysis methods	Summarized results
18/ B	Styve et al., 2022	CrisisLex (CrisisLexT6), Harvard Dataverse (Flood Tweet IDs)	Twitter	automated (LR, RF, CNN, ULMFiT, NER)	90 % of tweets were categorized as relevant by at least one classifier, however only 20-40 % were categorized as relevant by all classifiers. This shows that there are differences in the way individual tweets are classified and implies that the same tweets are not necessarily classified as relevant by each classifier, raising questions about the reliability of their predictions.
19/ A, B	Sufi and Khalil, 2022	none specific	Twitter	automated (NER, CNN, Getis-Ord Gi*, AR)	The purposed method can pinpoint a disaster location with 0.93 precision, 0.88 recall, 0.90 F1-Score, and 0.97 accuracy.
20/ C	Tavra et al., 2021	Forest fire, Croatia, 2017	multiple	manually (filtering, analysis and georeferencing)	The advantage of using multiple types of crowdsourcing data is that all age groups of citizens are covered and technical limitations can be overcome, such as a poor mobile connection. Data from the Crowdmap platform showed a low quantity but high spatial and temporal quality.
21/ D	Thom et al., 2015	Floods, Germany, 2013	Twitter	automated (ScatterBlogs)	decision makers confirm that analytical tools are helpful to manage the volume of data, identify eyewitness accounts and engage with citizens more broadly. However, it is important that analytical tools build on existing systems, language and identifiers. Classification, automatic event recognition, real-time capability, data protection and geo-referencing are considered particularly important. The aim should be to gather information from unmonitored areas as well as public opinion on response measures.
22/ C	Z. Wang and Ye, 2018	Hurricane Sandy, United States, 2012	Twitter	combination of automated + manually (annotation, LQ and Markov transition probability matrix)	The chosen approach of combinatorial evaluation of space, time, place and topic makes it possible to derive the focus of the population during the course of the disaster. In the period before and during the disaster, the main topics were infrastructure/utilities and weather/environment. During and after the disaster, donations/environments were discussed in particular.

Continued on next page

Table 2 – continued from previous page

No./ Focus	Reference	Disaster/ dataset	Platforms	Data analysis methods	Summarized results
23/ A, B	R.-Q. Wang et al., 2020	Hurricane harvey twitter dataset	Twitter	combination of automated + manually (NeuroNER, CV, CNN, ResNet v2)	The manual data check of automated classification shows significant advantages for credibility and reliability. By extracting information that official media may ignore, an orthogonal dimension can be opened up. In particular, information on street-level locations in SM can play a passive hotline that can support decision makers in the distribution of emergency response teams.
24/ A, B, C	Wu et al., 2022	Typhoon Lekima, China, 2019	Sina Weibo	automated (LDA, SnowNLP, OPTICS)	By linking multiple sources with SM data, deeper insights can be gained. In addition, negative sentiment extracted through sentiment analysis from microblogs can roughly show the spatial distribution of injuries and victims. The accuracy of extracting information about the actual victims at district level reached 31.8 % in the analysis conducted 72 hours after the typhoon hit.
25/ A, B	Yang et al., 2022	Flooding, China, 2020	Sina Weibo	automated (HanLP, NER)	One cannot get more disaster information by only paying attention to the posts with uploaded location tags, as there are few or no SM in some severely affected areas. The social network can be effectively combined with remote sensing image data and can help to gain more disaster information, such as assessing the disaster situation in different areas and analyzing the spatial distribution of people paying attention to flooded areas. By effectively combining data from multiple sources, the advantages of different data sources can be better utilized, contributing to a comprehensive description of disaster progression.

Continued on next page

Table 2 – continued from previous page

No./ Focus	*Reference*	*Disaster/ dataset*	*Platforms*	*Data analysis methods*	*Summarized results*
26/ A, B, C	Zhang et al., 2020	Hurricane Harvey, United States, 2017	Twitter	combination of automated + manually (SocialDISC: WSD, BiLSTM, SLPA)	SM analysis should (a) filter out irrelevant data, (b) sort information based on sub-events, needs, or user-defined categorizations, (c) reflect societal reactions to the impact of the disaster, (d) provide concise and easy-to-read results, and (e) ensure timely data processing. Analyzing the impact of Hurricane Harvey revealed that the disruption of housing and goods had the greatest impact on people's lives. In particular, tweets about relief and rescue were the most frequent among the nine categories and showed the most positive emotional signals. Information on societal impact, differences between expectations and realities, can be used in conjunction with other data as a reference for evaluating mitigation and preparedness plans for future disasters. Societal impact information derived from SM posts has the potential to improve disaster-specific models and assessment tools, contributing to increased community resilience.
27/ A, B	Zou et al., 2021	CrisisMMD and RE-SOURCE #9 (Cri-sisNLP)	Twitter	automated (adjusted VGG16, FastText model, Multimodal Fusion)	The presented multimodal approach consisting of deep learning method, FastText framework and data fusion model shows a better performance than the unimodal approach with 87.6 % and shows a better accuracy than the existing multimodal approach.

Abbreviations:
2PG: 2PhaseGeo, 2PGT: 2PhaseGeoTime, AC: Agglomerative Clustering, AGDenseNet: Attention Guided DenseNet, AR: automated regression, GIS: geographic information system, BERT: Bidirectional Encoder Representations from Transformers, BiLSTM: bidirectional Long Short-Term Memory, CNN: convolutional neural network, CNTM: Conditional Neural Turing Machine, CRF: Conditional Random Field, CV: Computer Vision, DDM: Damage Detection Map, DenseNet: Dense Convolutional Network, DT: Decision Tree, GCAM: Gradient-weighted Class Activation Mapping, IDW: Inverse Distance Weighted, KNN: KNeighbor, LDA: Latent Dirichlet Allocation, LQ: Location Quotient, LR: Logistic Regression, MLR: multiple linear regression, MNN: Multiple Neutral Network, NB: Na¨ıve Bayes, NER: Named Entity Recognition, NLP: Natural Language Processing, NN: neural network, OPTICS: Ordering Point to Identify the Cluster Structure, ResNet: residual neural networks, RF: Random Forests, SLPA: Uncovering Overlapping Communities in Social Networks via A Speaker-listener Interaction Dynamic Process, SocialDISC: semiautomated SM analytics approach for social sensing of Disaster Impacts and Societal Considerations, SOM: Self-Organizing Maps, SVM: Support Vector Machine, ULMFiT: Universal Language Model Fine-tuning, VGG: Visual Geometry Group, VOST: Virtual Operations Support Team, WEKA: Waikato Environment for Knowledge Analysis, WSD: word sense disambiguation, XGBOOST: Extreme Gradient Boosting, YOLO: You Only Look Once, Py: Python
Focus:
A: development/presentation of tool/framework, B: performance comparison of tools, C: gaining knowledge from SM (data insights), D: usefulness for decision makers

REFERENCES

Abid, S. K., Sulaiman, N., Chan, S. W., Nazir, U., Abid, M., Han, H., Ariza-Montes, A., & Vega-Muñoz, A. (2021). Toward an integrated disaster management approach: How artificial intelligence can boost disaster management. *Sustainability*, *13*(22), 12560. https://doi.org/10.3390/su132212560

Acikara, T., Xia, B., Yigitcanlar, T., & Hon, C. (2023). Contribution of social media analytics to disaster response effectiveness: A systematic review of the literature. *Sustainability*, *15*(11), 8860. https://doi.org/10.3390/su15118860

Alam, F., Alam, T., Ofli, F., & Imran, M. (2023). Robust Training of Social Media Image Classification Models. *IEEE Transactions on Computational Social Systems*, 1–20. https://doi.org/10.1109/tcss.2022.3230839

Asif, A., Khatoon, S., Hasan, M. M., Alshamari, M. A., Abdou, S., Elsayed, K. M., & Rashwan, M. (2021). Automatic analysis of social media images to identify disaster type and infer appropriate emergency response. *Journal of Big Data*, *8*(1). https://doi.org/10.1186/s40537-021-00471-5

Avvenuti, M., Del Vigna, F., Cresci, S., Marchetti, A., & Tesconi, M. (2015). Pulling Information from social media in the aftermath of unpredictable disasters. *2015 2nd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. https://doi.org/10.1109/ict-dm.2015.7402058

Dong, Z. S., Meng, L., Christenson, L., & Fulton, L. (2021). Social media information sharing for natural disaster response. *Natural Hazards*, *107*(3), 2077–2104. https://doi.org/10.1007/s11069-021-04528-9

Fan, C., Wu, F., & Mostafavi, A. (2020). A Hybrid Machine Learning Pipeline for Automated Mapping of Events and Locations From Social Media in Disasters. *IEEE Access*, *8*, 10478–10490. https://doi.org/10.1109/access.2020.2965550

Fathi, R., & Fiedrich, F. (2022). Social Media Analytics by Virtual Operations Support Teams in disaster management: Situational awareness and actionable information for decision-makers. *Frontiers in Earth Science*, *10*. https://doi.org/10.3389/feart.2022.941803

Fromm, J., Eyilmez, K., Baßfeld, M., Majchrzak, T. A., & Stieglitz, S. (2021). Social Media Data in an Augmented Reality System for Situation Awareness Support in Emergency Control Rooms. *Information Systems Frontiers*, *25*(1), 303–326. https://doi.org/10.1007/s10796-020-10101-9

Ghermandi, A., & Sinclair, M. (2019). Passive crowdsourcing of social media in environmental research: A systematic map. *Global Environmental Change*, *55*, 36–47. https://doi.org/https://doi.org/10.1016/j.gloenvcha.2019.02.003

Jixian, L., Gang, A., Zhihao, S., & Xiaoqiang, S. (2022). Social Media Multimodal Information Analysis based on the BiLSTM-Attention-CNN-XGBoost Ensemble Neural Network. *International Journal of Advanced Computer Science and Applications*, *13*(12). https://doi.org/10.14569/ijacsa.2022.0131215

Kankanamge, N., Yigitcanlar, T., Goonetilleke, A., & Kamruzzaman, M. (2020). Determining disaster severity through social media analysis: Testing the methodology with South East Queensland Flood tweets. *International Journal of Disaster Risk Reduction*, *42*, 101360. https://doi.org/10.1016/j.ijdrr.2019.101360

Karimiziarani, M. (2023). Social Media Analytics in Disaster Response: A Comprehensive Review. https://doi.org/10.48550/ARXIV.2307.04046

Kondraganti, A. (2021). Big Data Analytics in Humanitarian and Disaster Operations: A Systematic Review. https://doi.org/10.48550/ARXIV.2108.09800

Kuckartz, U. (2018). *Qualitative inhaltsanalyse. methoden, praxis, computerunterstu¨tzung* (4th ed.). Beltz Verlagsgruppe.

Lang, B., Dolan, R., Kemper, J., & Northey, G. (2020). Prosumers in times of crisis: Definition, archetypes and implications. *Journal of Service Management*, *32*(2), 176–189. https://doi.org/10.1108/josm-05-2020-0155

Lazreg, M. B., Goodwin, M., & Granmo, O.-C. (2019). An Iterative Information Retrieval Approach from Social Media in Crisis Situations. *2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. https://doi.org/10.1109/ict-dm47966.2019.9033008

Li, X., Zhang, H., Caragea, D., & Imran, M. (2018). Localizing and Quantifying Damage in Social Media Images. *CoRR*, *abs/1806.07378*.

Li, Z., Wang, C., Emrich, C. T., & Guo, D. (2017). A novel approach to leveraging social media for rapid flood mapping: A case study of the 2015 South Carolina floods. *Cartography and Geographic Information Science*, *45*(2), 97–110. https://doi.org/10.1080/15230406.2016.1271356

Lopez-Fuentes, L., Farasin, A., Zaffaroni, M., Skinnemoen, H., & Garza, P. (2020). Deep Learning Models for Road Passability Detection during Flood Events Using Social Media Data. *Applied Sciences*, *10*(24), 8783. https://doi.org/10.3390/app10248783

Luna, S., & Pennock, M. J. (2018). Social media applications and emergency management: A literature review and research agenda. *International Journal of Disaster Risk Reduction*, *28*, 565–577. https://doi.org/10.1016/j.ijdrr.2018.01.006

Mao, H., Thakur, G., Sparks, K., Sanyal, J., & Bhaduri, B. (2018). Mapping near-real-time power outages from social media. *International Journal of Digital Earth*, *12*(11), 1285–1299. https://doi.org/10.1080/17538947.2018.1535000

Moher, D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine*, *151*(4), 264. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Ning, H., Li, Z., Hodgson, M. E., & Wang, C. ((2020). Prototyping a Social Media Flooding Photo Screening System Based on Deep Learning. *ISPRS International Journal of Geo-Information*, *9*(2), 104. https://doi.org/10.3390/ijgi9020104

Pastor-Galindo, J., Nespoli, P., Gomez Marmol, F., & Martinez Perez, G. (2020). The not yet exploited goldmine of osint: Opportunities, open challenges and future trends. *IEEE Access*, *8*, 10282–10304. https://doi.org/10.1109/access.2020.2965257

Pereira, J., Monteiro, J., Silva, J., Estima, J., & Martins, B. (2020). Assessing flood severity from crowdsourced social media photos with deep neural networks. *Multimedia Tools and Applications*, *79*(35–36), 26197–26223. https://doi.org/10.1007/s11042-020-09196-8

Pohl, D., Bouchachia, A., & Hellwagner, H. (2013). Social media for crisis management: Clustering approaches for sub-event detection. *Multimedia Tools and Applications*, *74*(11), 3901–3932. https://doi.org/10.1007/s11042-013-1804-2

Styve, L., Navarra, C., Petersen, J. M., Neset, T.-S., & Vrotsou, K. (2022). A Visual Analytics Pipeline for the Identification and Exploration of Extreme Weather Events from Social Media Data. *Climate*, *10*(11), 174. https://doi.org/10.3390/cli10110174

Sufi, F. K., & Khalil, I. (2022). Automated Disaster Monitoring From Social Media Posts Using AI-Based Location Intelligence and Sentiment Analysis. *IEEE Transactions on Computational Social Systems*, 1–11. https://doi.org/10.1109/tcss.2022.3157142

Tavra, M., Racetin, I., & Perosˇ, J. (2021). The role of crowdsourcing and social media in crisis mapping: A case study of a wildfire reaching Croatian City of Split. *Geoenvironmental Disasters*, *8*(1). https://doi.org/10.1186/s40677-021-00181-3

Thom, D., Kruger, R., Ertl, T., Bechstedt, U., Platz, A., Zisgen, J., & Volland, B. (2015). Can twitter really save your life? a case study of visual social media analytics for situation awareness. *2015 IEEE Pacific Visualization Symposium (PacificVis)*. https://doi.org/10.1109/pacificvis.2015.7156376

von Elm, E., Schreiber, G., & Haupt, C. C. (2019). Methodische Anleitung fu¨r Scoping Reviews (JBI-Methodologie). *Zeitschrift fu¨r Evidenz, Fortbildung und Qualita¨t im Gesundheitswesen*, *143*, 1–7. https://doi.org/10.1016/j.zefq.2019.05.004

Wang, R.-Q., Hu, Y., Zhou, Z., & Yang, K. (2020). Tracking Flooding Phase Transitions and Establishing a Passive Hotline With AI-Enabled Social Media Data. *IEEE Access*, *8*, 103395–103404. https://doi.org/10.1109/access.2020.2994187

Wang, Z., & Ye, X. (2017). Social media analytics for natural disaster management. *International Journal of Geographical Information Science*, *32*(1), 49–72. https://doi.org/10.1080/13658816.2017.1367003

Wang, Z., & Ye, X. (2018). Space, time, and situational awareness in natural hazards: A case study of Hurricane Sandy with social media data. *Cartography and Geographic Information Science*, *46*(4), 334–346. https://doi.org/10.1080/15230406.2018.1483740

Wu, K., Wu, J., & Li, Y. (2022). Mining typhoon victim information based on multi-source data fusion using social media data in China: A case study of the 2019 Super Typhoon Lekima. *Geomatics, Natural Hazards and Risk*, *13*(1), 1087–1105. https://doi.org/10.1080/19475705.2022.2064774

Yang, T., Xie, J., Li, G., Zhang, L., Mou, N., Wang, H., Zhang, X., & Wang, X. (2022). Extracting Disaster-Related Location Information through Social Media to Assist Remote Sensing for Disaster Analysis: The Case of the Flood Disaster in the Yangtze River Basin in China in 2020. *Remote Sensing*, *14*(5), 1199. https://doi.org/10.3390/rs14051199

Zhang, C., Yao, W., Yang, Y., Huang, R., & Mostafavi, A. (2020). Semiautomated social media analytics for sensing societal impacts due to community disruptions during disasters. *Computer-Aided Civil and Infrastructure Engineering*, *35*(12), 1331–1348. https://doi.org/10.1111/mice.12576

Zou, Z., Gan, H., Huang, Q., Cai, T., & Cao, K. (2021). Disaster Image Classification by Fusing Multimodal Social Media Data. *ISPRS International Journal of Geo-Information*, *10*(10), 636. https://doi.org/10.3390/ijgi10100636

Utilizing Social Media and Crowdsourcing for Crisis Management and Communication: proposing an Integrative Framework

Kees Boersma
Department of Organization Sciences
Faculty of Social Sciences and
Faculty of Science
Vrije Universiteit Amsterdam
f.k.boersma@vu.nl

Nathan Clark
Department of Organization Sciences
Faculty of Social Sciences
Vrije Universiteit Amsterdam
n.e.clark@vu.nl

Robert Larruina
Department of Organization Sciences
Faculty of Social Sciences
Vrije Universiteit Amsterdam
r.l.larruina@vu.nl

Emmeline Roeloffs
Department of Organization Sciences
Faculty of Social Sciences and the
Netherlands Red Cross
Vrije Universiteit Amsterdam
e.n.roeloffs@vu.nl

ABSTRACT

Utilizing social media and crowdsourcing (SMCS) in crisis management is crucial for improving communication and citizen engagement. However, practitioners working in crisis management organizations often face challenges when translating the outcomes of scientific research into actionable knowledge. To aid disaster management organizations, a comprehensive set of resources including practical tools, guidelines, and use examples are necessary. This will help response organizations to integrate the different components of SMCS coherently and consistently in their organizations and to promote disaster resilience for different users. An integrative Framework, developed by the authors as part of the EU-funded Horizon 2020 project, LINKS – *strengthening links between technologies and society for European disaster resilience* , serves such resources. This paper presents how this Framework came into being and describes its main components. It then provides a case study from the Netherlands, focusing on the scenario of industrial, chemical hazards, to illustrate how the Framework was actually used.

Keywords

Social media, crowdsourcing, crisis management, disaster resilience, integrated framework.

INTRODUCTION

Beginning in the early 2000s, professional first responders, citizens, and local community members started to use social media platforms during crises to obtain information and assistance. Since that time, the use of social media platforms have been studied, and many empirical studies gave evidence to th e benefits for improved crisis management and crisis communication (Vieweg et al., 2010; Meier, 2012; Burgess et al., 2013; Hughes & Tapia, 2015). Indeed the integration of social media and crowdsourcing (SMCS) in crisis management has become increasingly important in recent years (Imran et al., 2018; Reuter & Kaufhold, 2018; Lovari, 2020; Haupt, 2021; Mavrodieva & Shaw, 2021; Harrison et al., 2024). This is due to its potentials for enhancing citizen engagement and improving communication during disaster situations.

Researchers have found that the use of social media in crisis management can greatly enhance disaster resilience. The usage of social media platforms can help with the quick collection and sharing of crucial information, tapping into the *collective intelligence* (Bruns et al., 2012, Palen et al., 2009; Palen & Hughes, 2018). Social media can also provide a platform for citizens to share information, mobilize resources, and coordinate activities (Reuter & Kaufhold, 2018; Reuter et al., 2018), and facilitate collaboration and coordination between disaster management organizations and citizens – to which we refer with the term crowdsourcing. The platform Ready2Help, developed and utilized by the Netherlands Red Cross provides a perfect example (Schmidt et al., 2 018; Boersma et al., 2019. This initiative utilizes an app to activate members from an extensive network of nearly 100,000 volunteers during emergencies.

However, practitioners working in the field of crisis and disaster often have a hard time translating the outcomes of scientific research into actionable knowledge. This paper focuses on how SMCS can be adopted by first responder organizations and argues for organizational change toward disaster resilience in a Dutch case study related to industrial (chemical) hazard scenarios. The main objective is to help disaster management organizations to improve crisis communication and citizen engagement. To this aim, this paper proposes an integrative Framework for implementing SMCS into crisis management, which can address the challenges that come with the use of social media in disaster situations. By bringing together different components required for successful SMCS utilization, the Framework can help disaster management organizations to navigate the complex landscape of social media communication and address challenges such as information accessibility, credibility and relevance, and misinformation. Furthermore, it can help ensure that all populations, including vulnerable ones, have equal access to timely and accurate information during disasters.

In the remaining part of this paper we would like to propose an *integrative Framework* that could support (research into) organizational, institutional change towards disaster resilience by enhancing knowledge for improved communication and citizen engagement. We will present a case study from the Netherlands to show how the Framework can be utilized. The case focusses on three organizations and their stakeholders: the public organization *Safety Region*[1] *of South Limburg*, responsible for managing emergencies and communicating risks to the community; *Sitech Services*, a technical service provider that offers advice and support to companies on emergency services at the *Chemelot site*, an industrial multi-user site and research campus that houses sixty-four industrial – mainly chemical - plants and about one hundred and fifty organizations, with over eight thousand employees.

SOCIAL MEDIA AND CROWDSOURCING: OPPORTUNITES AND CONSTRAINTS

Social media play an important role in disaster community resilience by enabling crisis communication, encouraging citizen involvement, and giving access to key information. It may give useful real-time data and information from many sources, boosting perception of the situation on the ground and helping disaster management organizations to respond more effectively (Apuke et al., 2018). Social media assist emergency response groups in communicating with the public, allowing for real-time updates, early warning systems, and the transmission of crucial information during a crisis. It can also help them to provide valuable data and information to other emergency management agencies to make informed decisions.

Social media has become an essential tool in crisis management for disaster management organizations and stakeholders, helping to engage with citizens, improve communication, increase public awareness, and ultimately save lives. SMCS has a great potential in engaging citizens in disaster management activities and to provide insights into integrating citizen-generated content for a more accurate response. Effective communication

[1] In the Netherlands, a Safety Region (Dutch: Veiligheidsregio) is a public body whose task is to facilitate regional cooperation in dealing with crises, disasters.

strategies that consider diverse needs and perspectives are necessary, particularly for vulnerable populations. SMCS also demonstrated importance in successful crisis response and recovery by facilitating *social connections*, which is especially important since social media may be used to organize response activities, mobilize resources, and offer real-time information to impacted people (Bukar et al., 2022).

These and many more studies have provided evidence that using social media in crisis management may significantly improve disaster resilience, crisis communication, and response efforts. *However*, incorporating social media communication strategies into crisis management has to overcome a number of obstacles.

First of all, there is a *practical* constraint as most response organizations lack the recourse and the organizational routines to integrate SMCS into their day-to-day (operational and strategic) practices (Boersma et al., 2016; Nielsen et al., 2023). If SMCS tools are not carefully integrated in crisis management plans, protocols and organizational routines, its effective and efficient use during crises cannot be ensured. In practice, the implementation of SMCS in organizations remains a struggle.

Secondly, while studies have shown the usefulness of SMCS for engaging citizens in crisis management and communication, they also revealed the *exclusion* of disadvantaged groups from essential information because of issues with information accessibility, such as the digital divide and language barriers. This exclusion can lead to a lack of timely and accurate information reaching those who need it the most, perhaps exacerbating the effect of a crisis (Lovari et al., 2020; Pérez-Escolar & Canet, 2023).

Thirdly, the rapid diffusion of information on social media platforms can lead to the *spread of disinformation*, which we have in particular seen during the COVID-19 pandemic, and which can have serious consequences during times of crisis (Fernandez and Alani, 2018; Burel et al., 2021). Misinformation travels quickly and widely, undermining public faith in official information sources and creating p ublic doubt. This has the potential to aggravate the problem and make reputation management more difficult (Roshan et al., 2016). Therefore, crisis communication plans need to be designed to identify and address misinformation, while also ensuring that accurate and relevant information is disseminated promptly and in an inclusive way.

A way forward to overcome such challenges is to develop and use an integrative Framework that can help stakeholders working in disaster management organization to utilize existing insights and focus on what's important when considering the application of SMCS in disasters. Furthermore, it can help in strategic thinking especially in the planning phases, and guides disaster management stakeholders through existing knowledge based on the objectives and needs of their organization.

METHODOLOGY

The data collected for this paper has been part of a project, called LINKS "Strengthening links between technologies and society for European disaster resilience", funded under the H2020 Research and Innovation Program, Disaster Resilient Societies (DRS). LINKS aimed to address the challenges of SCMC for crisis management and communication (Clark et al., 2024). An important outcome of the LINKS project is an integrative Framework that can serve as a knowledge resource for disaster management organizations a nd other relevant stakeholders to guide them in applying SMCS in disaster risk management. The guidance through the Framework is organized under two thematic areas: 1) engaging with citizens and 2) improving crisis and risk communication. The Framework has a user guidance, which provides simple access to the tools in the Framework through a dedicated website. A visualization of the Framework is presented in Figure 1, and detailed information is provided at the LINKS website: https://links-project.eu/links-Framework/.

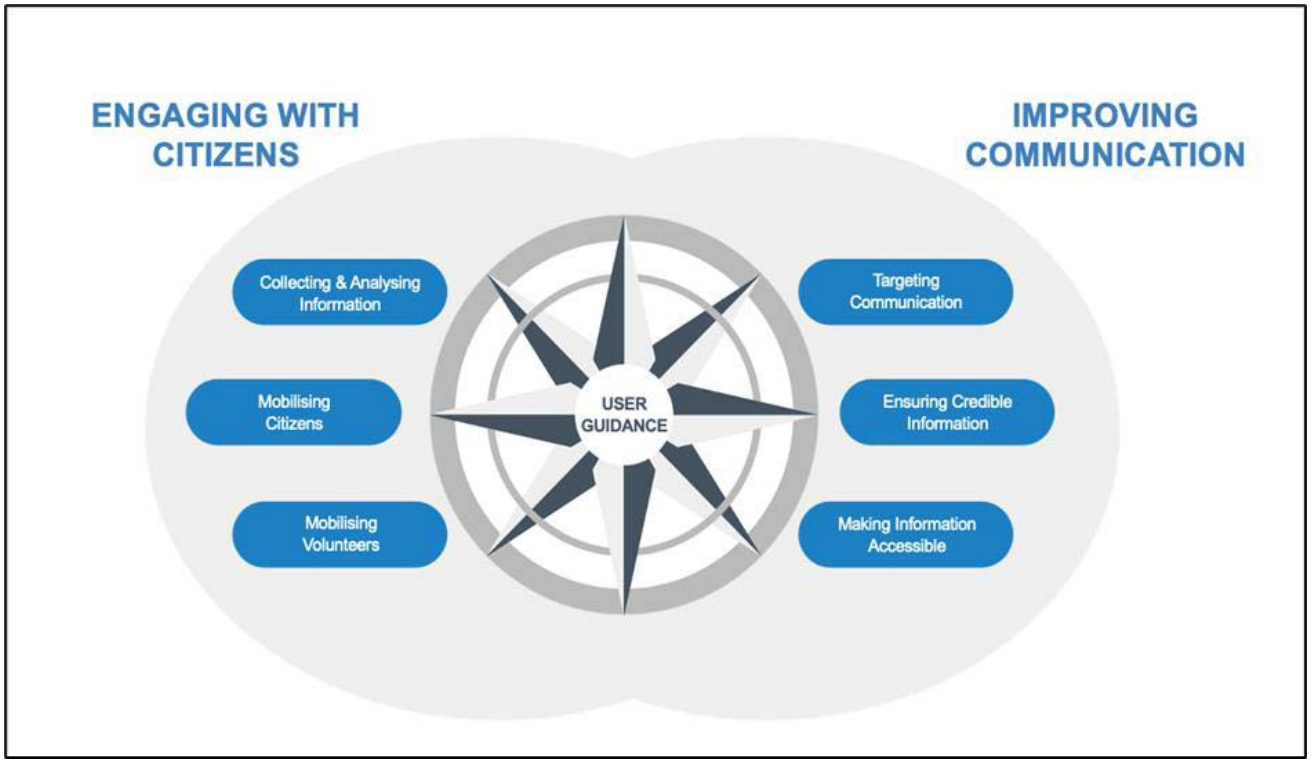


Figure 1. LINKS Framework adopted from Fonio et al., 2022

The Framework has been developed via an iterative process and based on a *participatory research approach* (Barreteau et al., 2013; Boersma et al., 2022) and on *action-research* (McNiff, 2013). The participatory dimension of our approach refers to the *co-creation* process that we adopted: in all phases of the research and the development of the Framework both academics and practitioners were involved from over 20 organizations. Usually it is hard to include citizens and community members in crisis management (Ferguson et al., 2018) but the Framework has been developed closely together with community representatives, labelled community *ambassadors*, whose role is explained in detail below. While the first few workshops were not organized in coordination with community representatives, over the course of several workshops the feedback from previous workshops and contacts in the field (both government and citizen contacts) have been used to better organize our workshops so they would be more appropriate for the audience and the needs. Currently this network is still used, in particular for communicating risk. On the long term this network has a big potential for active communication during a crisis.

In other words, the approach was action-oriented as the Framework was used as a tool to *intervene* in crisis management organizations to create organizational routines for the utilization of SMCS. The Framework was developed though different stages to do justice to both the participatory and the action -oriented dimension of the research.

In the first phase, we conducted 67 semi-structured interviews and an online survey with 284 responses across Europe to gain insights into the experiences, effective practices, and requirements of disaster management organizations and stakeholders regarding the use of SMCS for disaster management purposes. It was a vital part of our *participatory approach*. Specific themes were examined, such as decision-making protocols, sensitivity to vulnerability, reliable information, education, and technical aspects related to SMCS tools and processes used by disaster management organizations. The results were used to develop a first version of the Framework. Via five case studies related to specific hazards – drought, flooding, terrorist attacks, earthquakes and industrial incidents - we identified needs and gaps. The cases were contextualized in four European countries: the Netherlands, Denmark, Germany and Italy.

The second phase of the research was *action-oriented*: the Framework was implemented at case-level through workshops with practitioners (and other relevant stakeholders) to enhance crisis management (both response and preparation) and communication. Practitioners were asked to adopt SMCS in their crisis and communication plans using specific products embedded in the Framework, and also assisted with the development and updates of the products. Some of the most important issues we had to overcome organizing the workshops were:

- It was challenging to organize a workshop for shopkeepers, the general response was that they did not have the time to attend. We needed to adjust our workshop format for this group .
- We struggled to organize the workshop for civilians from the village Limbricht, which is located approximately 7.5 km away from the Chemelot site, citizens stated that they did not perceive any real danger for their town. We ended up organizing this workshop late in the evening on a non-holiday day.
- For those in the Lindenheuvel, directly neighboring the Chemelot site, the civilians had already been suspected to multiple research studies. Organizing the workshop in the Lindenheuvel required some mediation between other active (project) groups in the neighborhoods. This was done in order to prevent overloading the citizens any further.
- There were citizens who did not like hearing that they could contribute to preparedness as they felt that this was the responsibility of the formal authorities and the disaster management organizations. We heard reactions such as: "It is your job to keep us safe! Why does the responsibility have to land on our shoulders?". We were required to carefully navigate communicating safety in a way that did not feel leave the impression we were displacing responsibility.

Each workshop had its own focus group and included members of schools, hospitals, small businesses, shopkeepers. We also organized workshops with civilians in areas close to the Chemelot site (the neighborhoods of Lindenheuvel and Urmond) and one further away from the site (the neighborhood of Limbricht).

The workshops consisted of an informative component provided by professionals from the safety region of South - Limburg and from Sitech Services, a technical service provider. With information on the various roles of the disaster management organizations involved in the region during a chemical disaster, the upscaling of disaster scenarios and the actions that are taken by the disaster management organizations responsible during the various disaster phases.

Aside from an informative part the workshops also included an informal focus group discussion . The primary goal of the workshops was to ensure that citizens were made aware of the risks and that they know what to do in the event of a chemical incident. But perhaps most importantly, during the workshops we emphasized the importance of creating a plan and to practice this plan.

Participatory Action Research in the Dutch Case

During the initial stages of the project, we conducted qualitative and quantitative research into the communication practices of disaster management organizations active in the Safety Region of South-Limburg. In doing so we identified SMCS gaps in the communication practices of several disaster management organizations in the region. As an example, we found that the Safety Region utilizes the SMCS tool Obi4Wan. Obi4Wan is a very useful tool for data analysis and for monitoring social media activity. The inf ormation gathered through Obi4Wan is collected to make informed decisions during crises and for education and training purposes within the organization. The Safety Region of South Limburg uses its social media platforms to post about safety related information before and during crises. This information is often quite general, as this information is intended for a very large audience. The Safety Region of South Limburg does not have a tool which they can use to directly target specific communities or groups with risk and crisis communication. Nor is there a direct line of communication between the Safety Region and the community.

The participants in the meetings argued that organizing them around a certain *target group* or stakeholder (see Figure 2) would have the best effect if the participants could relate to a specific topic. In order to shape such target groups we reached out to the various organizations that belonged to a specific group or profession and left it up to the organizations to decide whom they chose to send as representative in what we called *community-based workshops*. To give an idea, for schools directors of several schools, members of school management, representatives of the school district (board) and representatives of the school's BHV team (bedrijfshulpverleening; the emergency response team of the school) represented educational professiona ls working in the near proximity of Chemelot. They brought in vital information about how schools (the school board, teachers and children) appreciated the industrial risks but also they could bring in information about the social fabric of the school's community that is useful for crisis management organizations.

The community-based workshops included: a discussion about the (formal) responsibilities of the disaster management organizations and governmental parties involved in preserving the safety of those at risk in the event of a chemical incident; an overview of the various actions disaster management organizations and governmental parties conduct in the preparation for and management of chemical incidents; a walkthrough of the safety protocols (what civilians are expected to do); a discussion on how civilians can improve their preparation and how disaster management organizations such as the Safety Region and Sitech can improve their communication in order to more directly target the needs of these groups in the communicative form that works best for them.

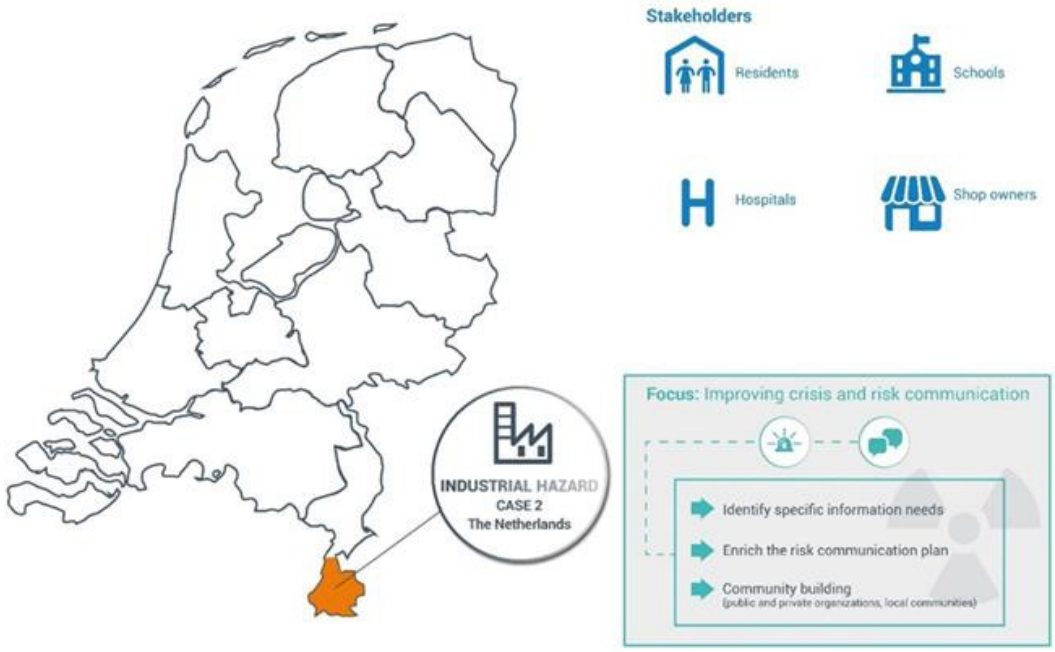


Figure 2. Case Area (from Clark et al., 2022)

Shaping an Ambassadors Network

The first community-based workshop was held in the municipality of Sittard-Geleen, neighboring the industrial site Chemelot, with the target group of *healthcare organizations*. This community-based workshop had eighteen participants, of which seven were healthcare professionals and eleven non -healthcare related professionals with a vested interest. This workshop included directors and healthcare professionals from various hospitals, healthcare practices and care homes from the surrounding municipalities. This workshop was primarily aimed at improving the risk awareness of the healthcare professionals, of making them aware that (chemical) accidents can happen and that proper preparation is therefore crucial, especially when there are so many moving parts. For example, staff on the road (on the way to clients or on the road with ambulances), staff at the homes of clients (care personnel), and staff at hospitals operating on patients (in the ICU).

The second targeted group was *schools* in from the various municipalities neighboring Chemelot (Beek, Stein, Sittard-Geleen). The workshop targeted school directors, school management, and school district representatives. The workshop had a total of thirty attendees, twenty-two were directors and upper management who represented the schools and board of education from the various municipalities. Of the remaining participants, five were organizers from the research team, and three were experts on risk and crisis communication (from the Fire Brigade, Sitech, Chemelot, the local government). This workshop had a similar goal of communicating the risks of a chemical incident, assessing the needs of the participants, providing an overview of the responsibilities of various parties and discussing the importance of creating a disaster plan. School directors, school management, and school district representatives appreciated the open discussion format of the community -based workshop, this provided the opportunity to exchange ideas, concerns and hurdles with us but also with one another.

The involvement of active representatives in the community-based workshops resulted in a network of *ambassadors* that helped us to *co-create* both the content and the implementation of the Framework and contributed to the understanding of the gaps in crisis communication and citizen engagement. The ambassadors can be seen as key figures who acted as 'boundary spanners' (Neal et al., 2022). Because the ambassadors were member of specific target groups *and* engaged in the risk awareness they were able to mobilize different stakeholders, such as healthcare institutions, schools (primary and secondary), and shop owners. The workshops brought together representatives of stakeholder groups that were identified a s priority target groups in the initial phase: health institutions, shop owners, and educational institutions. The meetings provided insight into the information needs of various stakeholders, investing in a sustainable community of professionals and resid ents, and supporting entrepreneurs, companies, and institutions in their preparations for a potential chemical incident.

RESULTS

Throughout the various workshops, we learnt about the vast number of online groups such as neighborhood WhatsApp groups or sport club WhatsApp groups. A single targeted safety message in one of these groups has the potential of being read by hundreds of individuals. The ambassadors advised us on how we can best target the groups or communities they help represent. They could give valuable input to the crisis management professionals present at the workshops on the type of preferred communication format (digital, print, word of mouth), langua ge and information is preferred. In addition, the ambassadors help share safety information about the chemical industry through their networks and help expand our existing network.

During the workshops the participants indicated that building and maintaining relationships with stakeholders is crucial, not only during emergencies (response phase) but also in 'normal' circumstances (preparation phase). This seems to be an obvious result, but it was striking how difficult it was for the different stakeholder to find each other despite the good intentions. Social media platforms were seen as a promising tool for making new connections, but at the same time their rapid changing popularity makes it difficult for the users to fully utilize their potential and to integrate them in the first responders' crisis communication strategies.

The community-based workshops' participants agreed that understanding *how* information is shared and received by citizens and professionals is key and depends on the characteristics of a specific platform . Not knowing how to use them and how to make sense of the data creates confusion. For example, various publicly accessible social media sources are used by the Safety Region Limburg South to gauge information needs and sentiments, but there are doubts about how representative these sources are of how the inhabitants of the municipalities actually feel about hazards and risks. Communication between citizens increasingly occurs in non-public spaces, such as neighborhood apps or family groups, which can make it challenging to assess information needs and sentiment towards organizations.

The discussions revealed that adequate and precise crisis communication is challenging in case of a chemical spill. For example, recalling people back inside or instructing them to stay indoors during a chemical incident was identified as an area that needs more attention. The workshop participants emphasized the need for better communication systems and guidelines within healthcare organizations to ensure that critical information is disseminated effectively.

Finally, we found a need among the first responders to clarify the concept of *crowdsourcing*, as they struggle with finding the right tools to mobilize citizens. The lack of internal guidelines, organizational routines but also laws and regulations regarding social media and crowdsourcing during crises and disasters were seen as serious constraints.

Zooming in: Community-Based Workshops with Healthcare and Educational Professionals

The community-based workshops we organized with and for *healthcare professionals* identified several key conclusions regarding the preparedness of healthcare organizations for chemical incidents. Participants highlighted the need for more emphasis on self-reliance and preparedness of healthcare organizations and their staff, indicating a gap in current practices. They also expressed surprise that they were supposed to practice and prepare for chemical incidents, and they indicated the need for greater awareness and room for training to ensure that healthcare professionals are adequately prepared for such events.

The healthcare professionals appreciated being able to communicate their reservations and needs directly during the workshop. They pointed out that they would like more direct and targeted communication between disaster management organizations and the communication teams of the respective hospitals and healthcare practices. For example, questions raised by the workshop's participants were: "If we need to close the doors and windows, are we allowed to let our staff inside?", "If we turn of the ventilation or IC units won't function anymore.", "A lot of our care home workers are often on the road with the bike or car, how do we preserve their safety?" and – a returning question - "How do we make sure people stay indoors during an incident?"

They hoped for more targeted information during crises but also assistance in creating disaster plans for their own organizations. The workshop participants emphasized the need for incident training exercises for healthcare organizations and suggested that existing structures, such as hospital ventilation and organizational guidelines, should be re-evaluated and adapted for chemical incidents. These findings suggest that ongoing preparedness efforts are needed to ensure the safety of healthcare professionals and patients in the event of a chemical incident and that a comprehensive approach is required to address the identified gaps in current practices.

The community-basedworkshops with *educational professionals* identified several key conclusions related to the preparedness of educational systems for chemical incidents, highlighting the need for accessible and functional

windows, doors, grilles, and ventilation systems that can be quickly and easily closed off in the event of an incident. The workshop also recommended that schools have a pre-prepared (digital) message for parents, teachers, and children, as well as a meter card for chemical incidents and a well-coordinated communication plan. Additionally, the workshop emphasized the need for schools to establish a clearance program and practice different scenarios to ensure they are well-prepared to respond to any potentia l incidents.

Some of the school directors advised the crisis management professionals to practice exercises with their schools and others requested assistance with their disaster plans. The schools commented that they would appreciate more direct communication with the Safety Region in particular, they felt it would also be beneficial if there was a way to improve the communication between the schools themselves. It can be confusing for parents with children that go to different schools when not every school reacts the same way to incidents. For example, certain schools letting the children go while other schools keep the children inside for safety reasons. Examples of the kind of questions and issues for discussions during these meetings were: "What do we do if there are parents at the door, but we are required to keep students indoors for their safety?" and "We had an incident a while back where students from a neighboring district were allowed to go home, but we kept the students indoors. What do we tell the angry parents?"

The community-based workshops (guided by the implementation of the Framework) proved to be effective in contributing to the development of effective tools and resources for enhancing preparedness and response capabilities in the event of a chemical incident. By sharing practical perspectives, they contributed to the development of a more comprehensive and inclusive framework for engaging citizens in various crisis contexts.

DISCUSSION AND CONCLUSION

In this paper, we argued for the need to develop a coherent set of tools and guidelines – an integrated Framework - to successfully implement, utilize and further develop social media and crowdsourcing in crisis and disaster management. SMCS tools were seen as promising, but only if they serve the needs of local communities and under the condition that they could be easily integrated in the first responders organizational routines. And while there is a lot of knowledge on both the benefits and the challenges of SMCS in crisis context, both academics and practitioners struggle to integrate the knowledge coherently and consistently. By providing a framework based on co-creation activities, we believe the use of SMCS can be more successful. Next, we introduced the Framework developed in the European project LINKS and its implementation in a case study on industrial, chemical hazards.

The case study illustrates how the Framework helped to identify key components for effective SMCS integration in crisis management. However, continuous iteration is necessary for a successful use and integration of SMCS in crisis management and communication practices. The iterative process and the participatory and action-oriented case study approach allowed refinement to address emerging challenges and specific needs. The community-based workshops served as a means to apply and refine SMCS integration. Local community ambassadors networks were vital to developing effective tools and resources for enhancing preparedness and response capabilities in specific contexts.

The case has also shown how the Framework can contribute to organizational change by enhancing preparedness and response capabilities for chemical incidents in healthcare and educational institutions. By integrating the feedback from community-based workshops professional first response organizations made a start with the development of effective practices for the integration of SMCS. Perhaps the most important lesson when it come sot the successful use of social media for crisis and risk communication is that it should be combined with or contain concrete ' action perspectives' and that face-to-face meetings with those who are supposed to follow up the messages are crucial.

Finally, the ambassador program is an example of how crisis management organizations can be actively engaged with a diverse range of community organizations, including educational institutions, healthcare institutions, and neighborhood support groups to enhance preparedness and response capabilities for chemical incidents in municipality neighboring a chemical hazardous site.

A Promising Outlook

The Framework's focus on local implementation has allowed for a tailored approach that considers local diversity, vulnerability, accessibility, and institutional learning processes, enhancing disaster resilience and coordination between authorities and citizens. The co-created nature of the Framework proposed in the paper has contributed to its legitimacy. Looking to the future, the Framework can contribute to development of organizational and institutional routines based on the experiences of both professional first responders and citizens. By sharing practical perspectives, it will continue to help shape the development of more comprehensive and effective

resources for engaging citizens and improving the use of SMCS, beyond the case of Chemelot.

We believe that using such integrative frameworks for SMCS based on co—creation approaches will result in better communication practices in disaster management and ultimately contributing to a safer and more resilient society.

ACKNOWLEDGMENTS

This work was supported by the LINKS Project that has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement No. 883490. We would like to thank the project's partners Safety Region of South Limburg, Sitech Services and Chemelot for providing access and for co-organizing the workshops.

REFERENCES

Apuke, O. D., & Tunca, E. A. (2018). Social media and crisis management: a review and analysis of existing studies. *LAÜ Sosyal Bilimler Dergisi, EUL Journal of Social Sciences*, 9(2), 199-215.

Barreteau, O., Bots, P., Daniell, K., Etienne, M., Perez, P., Barnaud, C., ... & Trebuil, G. (2013). 'Participatory approaches.' In: Edmonds, B, & Meyer, R. (Eds.), *Simulating Social Complexity* (197-234). Berlin: Springer,.

Boersma, F. K., Diks, D., Ferguson, J., & Wolbers, J. (2016). 'From reactive to proactive use of social media in emergency response: A critical discussion of the Twitcident Project'. In: Silvius, G. (Ed.). *Emergency and Disaster Management: Concepts, Methodologies, Tools, and Applications* (pp. 236-252). IGI Global.

Boersma, F. K., Ferguson, J., Groenewegen, P., Mulder, F., Schmidt, A., & Wolbers, J. (2019). 'Platform governance of self-organized initiatives in response to disasters'. In: Koppenjan, J., Termeer, K., & Karré, P. M. (Eds.). *Smart hybridity: Potentials and challenges of new governance arrangements* (31-42). Den Haag: Boom.

Boersma, F. K., Berg, R., Rijbroek, J., Ardai, P., Azarhoosh, F., Forozesh, F., ... & Bos, J. (2022). Exploring the potential of local stakeholders' involvement in crisis management. The living lab approach in a case study from Amsterdam. *International Journal of Disaster Risk Reduction*, 79, 103179.

Bruns, A., Burgess, J., Crawford, K., & Shaw, F. (2012). *Crisis Communication on Twitter in the 2011 South East Queensland Floods*. Brisbane: ARC Centre of Excellence for Creative Industries and Innovation.

Bukar, U. A., Jabar, M. A., Sidi, F., Nor, R. B., Abdullah, S., & Ishak, I. (2022). How social media crisis response and social interaction is helping people recover from Covid-19: an empirical investigation. *Journal of Computational Social Science*, 5, 781-809.

Burel, G., Farrell, T., Mensio, M., Khare, P., & Alani, H. (2020). 'Co-spread of misinformation and fact-checking content during the Covid-19 pandemic'. In: *Social Informatics: 12th International Conference Proceedings* (28-42). Cham: Springer.

Burgess, J., Bruns, A., Crawford, K., Finn, M., Monroy-Hernandez, A., & Palen, L. (2013). Social Media in Crisis Communication. *AoIR Selected Papers of Internet Research*, 3.

Clark, N., Fonio, C., Lüke, R., Bonati, S., Nardini, O., Graziani, F., Claessens, M., Rijkx, L., Andersen, N., Thayssen, J., Rammert, S., Hammachers, A., & Hingmann, N. (2022). First LINKS Case Report. Deliverable 6.4 of LINKS Project, funded by the EU Horizon 2020 Research and Innovation Programme (No. 883490). Retrieved from http://links-project.eu/deliverables/

Clark, N., Boersma, F. K., Bonati, S., Fonio, C., Gehlhar, S., Habig, T., Larruina, R., Lüke, R., Morelli, S., Bach Nielsen, A. Opromolla, A., Pazzi, V., & Raju, E. (2024). Exploring the impacts of social media and crowdsourcing on disaster resilience [final version 3]. *Open Research Europe*. https://open-research-europe.ec.europa.eu/articles/1-60/v3

Ferguson, J., Schmidt, A., & Boersma, F. K. (2018). Citizens in crisis and disaster management: Understanding barriers and opportunities for inclusion. *Journal of Contingencies and Crisis Management*, 26(3), 326-328.

Fernandez, M., & Alani, H. (2018). Online misinformation: Challenges and future directions. In : *Proceedings of the The Web Conference*, 595-602.

Fonio, C., Clark, N., Bonati, S., Lüke, R., Graziani, F., Habig, T., Nielsen, A., & Raju, E., (2022). First version of the LINKS Framework. Deliverable 5.3 of LINKS Project, funded by the EU Horizon 2020 Research and Innovation Programme (No. 883490). Retrieved from http://links-project.eu/deliverables/

Harrison, S. E., Tan, M. L., Prasanna, R., & Doyle, E. E. (2024). Social media use by multiple stakeholders and rightsholders during the Fort McMurray wildfires. In *Case Studies in Disaster Response* (pp. 213-236). Butterworth-Heinemann.

Haupt, B. (2021). The Use of Crisis Communication Strategies in Emergency Management. *Journal of Homeland Security and Emergency Management*, 18(2), 125-150.

Hughes, A. L., & Tapia, A. H. (2015). Social Media in Crisis: When Professional Responders Meet Digital Volunteers. *Journal of Homeland Security and Emergency Management*, 12(3), 679-706.

Imran, M., Meier, P. & Boersma, F. K. (2018). 'The use of social media for crisis management: a privacy by design approach', in: Boersma, F. K., & Fonio, C. (eds.) (2018). Big Data, Surveillance and Crisis Management. London: Routledge, pp. 19-37.

Lovari, A., & Bowen, S. A. (2020). Social media in disaster communication: A case study of strategies, barriers, and ethical implications. *Journal of Public Affairs*. 20(1), e1967.

Mavrodieva, A., & Shaw, R. (2021). 'Social Media in Disaster Management'. In: Shaw, R., Kakuchi, S., & Yamaji, M. (Eds.). *Media and Disaster Risk Reduction: Advances, Challenges and Potentials* (55-73). Singapore: Springer.

McNiff, J. (2013). *Action research: Principles and practice*. London: Routledge.

Meier, P. (2012). Crisis Mapping in Action: How open source software and global volunteer networks are changing the world, one map at a time, *Journal of Map & Geography Libraries*, 8(2), 89-100.

Neal, J. W., Neal, Z. P., & Brutzman, B. (2022). Defining brokers, intermediaries, and boundary spanners: a systematic review. *Evidence & Policy*, 18(1), 7-24.

Nielsen, A. B., Raju, E., Landwehr, D. R., Nicolaï, J. E., Patil, T. & Andersen, N. B. (2023). Leveraging Social Media and Crowdsourcing in Disaster Risk Management Processes in Europe. LINKS Project, funded by the EU Horizon 2020 Research and Innovation Programme (No. 883490).

Palen, L., Vieweg, S., Liu, S., & Hughes, A. (2009). Crisis in a Networked World: Features of computer-mediated communication in the April 16, 2007, Virginia Tech event. *Social Science Computer Review*, 27(4), 467-480.

Palen, L., & Hughes, A. L. (2018). 'Social media in disaster communication'. In: Rodriguez, H., Donner, W., & Trainor, J.E. (Eds.). *Handbook of disaster research* (497-518). New York: Springer.

Pérez-Escolar, M., & Canet, F. (2023). Research on vulnerable people and digital inclusion: toward a consolidated taxonomical Framework. *Universal Access in the Information Society*, 22(3), 1059-1072.

Reuter, C., Hughes, A. L., & Kaufhold, M. A. (2018). Social media in crisis management: An evaluation and analysis of crisis informatics research. *International Journalof Human-Computer Interaction*, 34(4), 280-294.

Reuter, C., & Kaufhold, M. A. (2018). Fifteen years of social media in emergencies: a retrospective review and future directions for crisis informatics. *Journal of Contingencies and Crisis Management*, 26(1), 41-57.

Roshan, M., Warren, M., & Carr, R. (2016). Understanding the use of social media by organisations for crisis communication. *Computers in Human Behavior*. 63, 350-361.

Schmidt, A., Wolbers, J., Ferguson, J., & Boersma, F. K. (2018). Are you Ready2Help? Conceptualizing the management of online and onsite volunteer convergence. *Journal of Contingencies and Crisis Management*, 26(3), 338-349.

Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010). 'Microblogging during two natural hazards events: what twitter may contribute to situational awareness' In: *Proceedings of the SIGCHI conference on human factors in computing systems*, 1079-1088.

Effectiveness of Emergency Response to Typhoons: Lessons Learned from Historical Typhoon Disasters in Zhuhai, China

Ting Liu
Joint International Research Laboratory of Catastrophe Simulation and Systemic Risk Governance, Beijing Normal University, Zhuhai 519087, China;
School of National Safety and Emergency Management, Beijing Normal University, Zhuhai 519087, China
202321099011@mail.bnu.edu.cn

Xiaoyong Ni*
Joint International Research Laboratory of Catastrophe Simulation and Systemic Risk Governance, Beijing Normal University, Zhuhai 519087, China;
School of National Safety and Emergency Management, Beijing Normal University, Zhuhai 519087, China
nixiaoyong@bnu.edu.cn

Jiawei Sun
Faculty of Arts and Sciences, Beijing Normal University, Zhuhai 519087, China
202211079130@mail.bnu.edu.cn

Ziyi Peng
Bay Area International Business School, Beijing Normal University, Zhuhai 519087, China
202211109007@mail.bnu.edu.cn

*Correspondence: nixiaoyong@bnu.edu.cn

ABSTRACT

Typhoons significantly impact socio-economics, varying in impact due to unique characteristics. However, limited studies explore why similar intensity typhoons affect regions differently. This study compares the emergency response measures of Zhuhai, Guangdong Province, China, in response to typhoon Hato (2017), Mangkhut (2018), and Higos (2020). These three similar typhoons offer an empirical opportunity to assess the effectiveness of emergency management. This research builds a timeline of Zhuhai's emergency response measures to the typhoons, collects Weibo data 72 hours before and after the typhoon landfall, performs keyword frequency analysis on contents and analyzes in detail the content of weibo postings, and changes in sentiment. Results show that timely and systematic emergency response measures may reduce disaster losses, which means emergency management is crucial in mitigating typhoon impacts, offering valuable lessons for other coastal cities.

Keywords

Typhoon disaster, emergency management, comparative analysis, data mining.

INTRODUCTION

Typhoon is an extremely destructive natural disaster whose impact spans the corners of globe (Gettelman et al., 2018). Worldwide, about 150 million people are affected by typhoons each year (Geiger et al., 2021). The Typhoon can trigger strong winds, heavy rainfall, storm surges and other disasters, posing a significant threat to human life safety and social property, and causing severe impact on industries such as agriculture, fisheries, and transportation. For example, Hurricane Katrina in 2005 caused severe flooding in areas such as New Orleans, which dislocated a large number of residents and resulted in more than 1,800 deaths and an estimated $125 billion in economic damages, also exposing the humanitarian crises (Guidotti, 2006). Similarly, Hurricane Haiyan in the Philippines in 2013 affected about 12.2 million people, causing almost 490,000 homes completely destroyed and

WiP Paper – Social Media for Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

more than 590,000 homes partially destroyed, resulting in economic losses estimated at PhP 424 billion , representing 3.7 percent of GDP that year in Philippines (Santos et al., 2016). In addition, Typhoon Haima in 2016 (Zhang et al., 2019) and Super Typhoon Mangosteen in 2018 (Choy et al., 2020) wreaked havoc in the Philippines and Southeast Asia, causing thousands of loss lives and devastating the local economy and social stability. These cases highlight the threat and danger of typhoons as an extreme weather phenomenon. Studies have shown that under global warming and climate change, the intensity and uncertainty of typhoons are continuing to increase, and the number and proportion of typhoons with high intensity will rise (Xu et al., 2020). Hence, there is an urgent need to attach greater importance to typhoon disasters and the corresponding emergency management measures.

Typhoon disaster emergency management encompasses organized, planned, and coordinated activities to mitigate the impact of typhoons on lives, property, and the socio-economy. Worldwide research focuses on risk assessment, early warning, emergency response, and post-disaster recovery. Risk assessment is crucial for comprehensive disaster mitigation and emergency management, helping prevent and reduce losses (Shiyuan et al., 2006). Early warning systems are essential, and researchers aim to improve prediction accuracy using modern meteorological and information technology (Huo et al., 2019; Jiang et al., 2018). Innovations like drone technology enhance monitoring, visualization, and analysis, strengthening meteorological and disaster forecasting collaboration (Chen et al., 2020). Multi-departmental and multi-level warning systems are established and improved, with various release channels to ensure timely warnings to the public. Post-typhoon reconstruction aims to restore normalcy, emphasizing environmental protection and sustainable resource use. It also promotes social psychological recovery and economic revitalization. Comprehensive measures, including enhanced early warning sy stems and improved policies, aim to address the profound impact of typhoons, ensuring a sustainable an d steady future for society (Sylwanowicz et al., 2018).

Typhoon disaster emergency management hinges on swift and effective emergency response. Scientific typhoon emergency plans delineate departmental responsibilities, coordinate resources, and ensure rapid, effective responses. A study retrieve the most similar historical cases, providing a reference for generating emergency plans for target typhoon trajectories (Li, 2022). Scientific evacuation plans (Sun et al., 2020) and the selection of emergency shelter locations (Anping, 2010) are directly related to personnel safety. A model based on typhoon track changes aids in emergency shelter allocation (Qin et al., 2019). Meanwhile, the reserve and demand forecasting of relief supplies are crucial aspects of emergency response (Pan et al., 2017; Uichanco, 2022). By predicting the real-time and future needs of affected areas, various types of relief supplies can be reasonably stocked to ensure timely and effective rescue operations after disasters occur. Additionally, the adequacy of household emergency preparedness has garnered attention. Through investigative research, scholars have found that enhancing household awareness and preparatory measures for typhoon events can effectively reduce disaster-related losses (Chan et al., 2019). The timeliness and effectiveness of post-disaster rescue are crucial for mitigating disaster losses and safeguarding people's lives (Santiago et al., 2016). The restoration of water networks (Wei et al., 2018), power grids (Liu et al., 2023), communications (G.C et al., 2019), and transportation (Wang et al., 2018) is essential for resuming normal social functions in affected areas. Researchers are dedicated to enhancing the disaster resistance of these infrastructures and studying efficient restoration solutions to expedite the return to normal social operations after disasters. The recovery of the healthcare system is a critical aspect of typhoon disaster emergency response. Scholars focus on ensuring adequate medical resources, timely emergency services, and optimizing the organizational structure of the healthcare system to meet medical needs during disasters (Lai et al., 2003; Su, 2013). In recent years, social media has demonstrated a significant role in crisis response (Liu et al., 2020), particularly in typhoon disaster emergency management (Athanasia & Stavros, 2015; Pollack et al., 2015). Its immediacy and interactivity make it a pivotal channel for information dissemination, aiding the public in quickly accessing warnings and emergency measures. Social media content analysis offers new insights and means for risk assessment and early warning. Furthermore, social media facilitates the coordination and integration of rescue efforts, enhancing rescue efficiency. For instance, studies have employed sentiment analysis through artificial neural networks to assess the urgency of relevant social media information, thereby enhancing the efficiency and effectiveness of emergency management (Ermino et al., 2022).

China, with its extensive coastlines and unique geography, is highly susceptible to typhoons (Wang et al., 2021), particularly in the major deltas of Tianjin, Shanghai, and Guangzhou (Yin et al., 2013). The southeast coast experiences frequent typhoons and significant damage, while rapid economic development and population density have increased economic losses from typhoon disasters (Li et al., 2016). As a coastal city in Guangdong Province, Zhuhai is located on the west bank of the Pearl River and is directly affected by the typhoon season. Historical records show that Zhuhai has endured numerous severe typhoons, posing constant challenges to urban management and disaster prevention and mitigation efforts. Differences in actual damage caused by similar-intensity typhoons are observed. This study compares three recent typhoons—Pigeon (2017), Mangkhut (2018), and Higos (2020)—with similar paths and intensity, all significantly impacting production and living in Zhuhai.

Using a comprehensive big data approach, this study analyzes news and social media data on Zhuhai's typhoon disaster emergency management. Through word frequency and categorical analysis, we aim to understand the evolution of Zhuhai's emergency response to similar-intensity typhoons. Results show Zhuhai's consistent optimization of emergency measures, which improves the city system's resilience to natural disasters. This study provides theoretical insights for disaster prevention, mitigation, and sustainable urban resilience.

OVERALL STRUCTRUE

This study (Figure 1) focuses on three recent typhoons hitting Zhuhai — Hato (2017), Mangkhut (2018) and Higos (2020) — with similar tracks and intensities. Data mining is used to collect relevant news articles and Weibo data tagged with 'Zhuhai' and 'typhoon', while web crawling retrieves Weibo during a 144-hour window, encompassing 72 hours before and 72 hours after each typhoon's landfall, forming a comprehensive disaster event dataset. Analyzing Weibo content via lexical frequency, word clouds and categorization, the study uncovers prominent public concerns and tracks changes in emotional sentiment throughout the warning-to-recovery cycle. It offers a new perspective on disaster management by vividly illustrating these fluctuations in a time series graph.

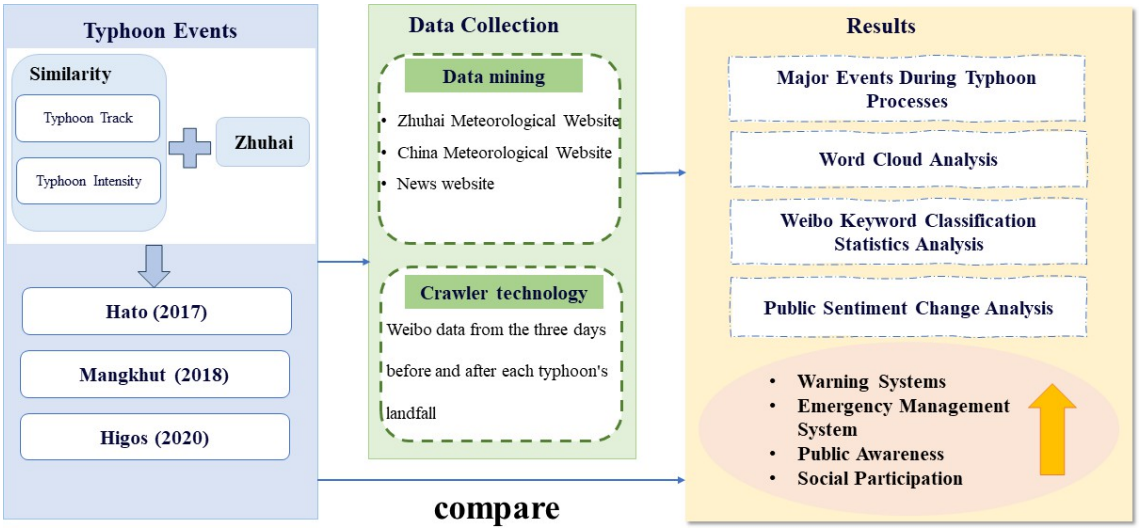


Figure 1. The overall structure of this paper

STUDY AREA AND TYPHOON EVENTS

Study Area

Zhuhai, situated in Guangdong Province's southwestern Pearl River Delta, is a central Pearl River Delta city and boasts the region's largest ocean area, most islands, and longest coastline. It uniquely connects to both Hong Kong and Macao via road. Covering 1,725 km² of land and 9,348 km² of sea within its territorial baseline, Zhuhai had 2.477 million residents by late 2022, with a 90.76% urbanization rate. Administered through Xiangzhou, Doumen, and Jinwan districts, it includes 15 counties, 10 streets, a nd 329 villages/communities (Figure 2). Zhuhai experiences a subtropical monsoon climate, prone to various catastrophic weather influences and emergencies, particularly typhoons in summer and autumn. Typhoons bring substantial unpredictability and complexity, inflicting wind, rain-induced flooding, and coastal storm surge disasters on Zhuhai. The city averages around 4 typhoons annually, with 1 severe impact and 5 instances of heavy rainfall, according to the Zhuhai Meteorological Bureau's "White Paper on Meteorological Public Services". (Su et al., 2022)"

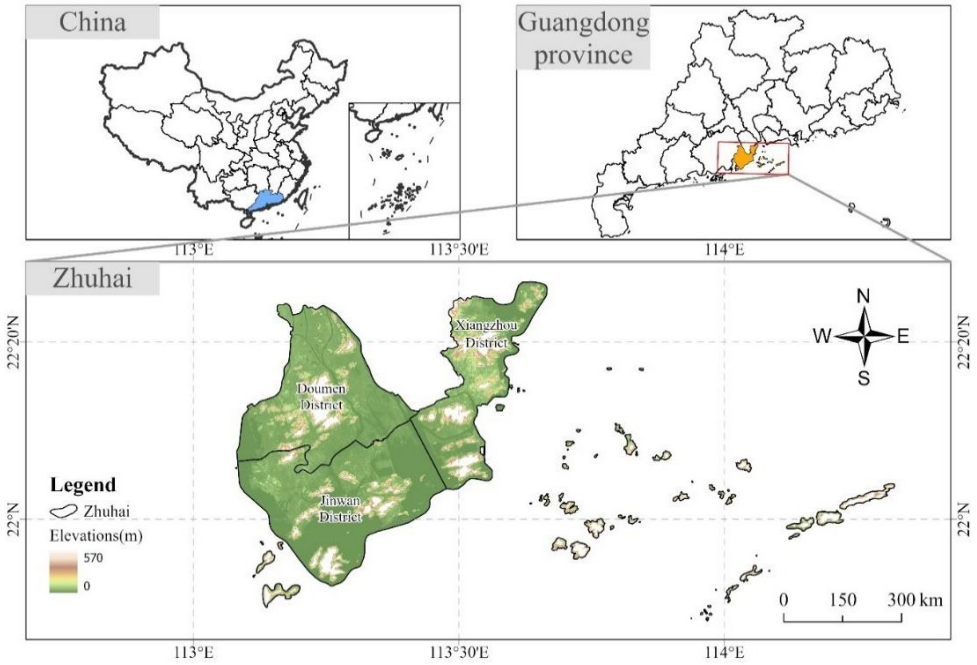


Figure 2. Geographic location of zhuhai

Typhoon Events

Using the 1949-2022 tropical cyclone dataset from the China Meteorological Administration (Lu et al., 2021; Ying et al., 2014), a 300 km radius around Zhuhai was analyzed for typhoon impact. Excluding the unnumbered typhoons, a total of 1540 typhoons are collected in the dataset, and finally 228 typhoons affecting Zhuhai are extracted, averaging 3.08 typhoons per year. Typhoons Hato (2017), Mangkhut (2018), and Higos (2020) are selected as the study objects (Table 1) based on their intensity, wind speeds, trajectories (Figure 3), and impact on Zhuhai.

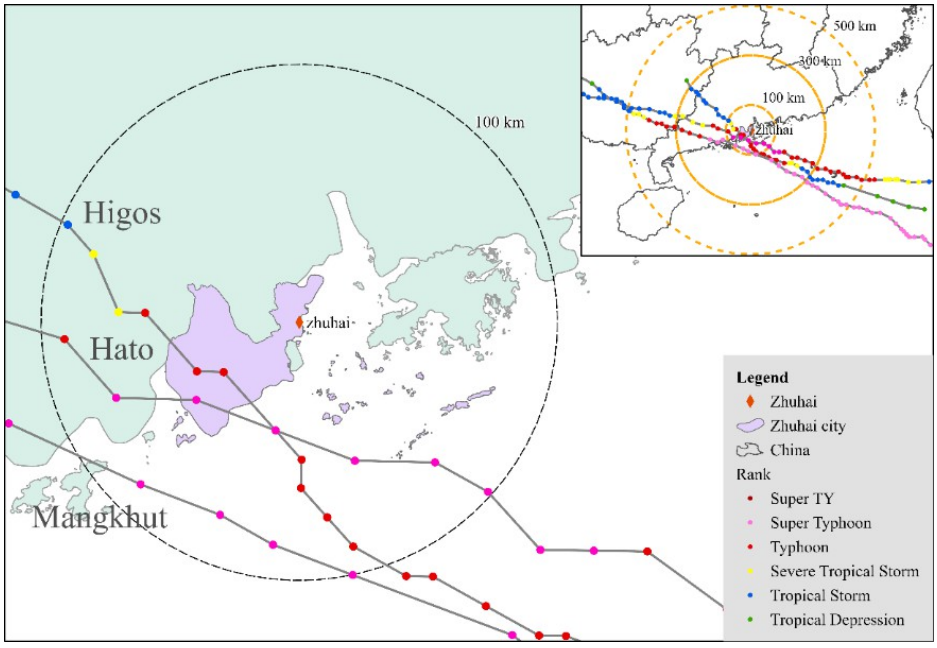


Figure 3. Trajectories of typhoon Hato, Mangkhut, Higos

Table 1. Disaster Information of Typhoons and its Impact on Guangdong Province

Typhoon	Hato	Mangkhut	Higos
Year	2017	2018	2020
Generation time(month/day)	8/20	9/7	8/16
Landing time(month/day)	8/23	9/16	8/19
Landing site	Zhuhai, Guangdong	Taishan, Guangdong	Zhuhai, Guangdong
Landing intensity	Super Typhoon	Super Typhoon	Typhoon
Maximum wind power (level) (wind speed, m/s)	14(48)	14(42)	12(33)
Affected areas (mainland China)	Guangdong, Guangxi, Yunnan, Hunan, Guizhou, Fujian	Guangdong, Guangxi, Hainan, Hunan, Yunnan, Guizhou	Guangdong, Guangxi
Impact on Guangdong Province			
Affected population(10,000 people)	142.4	306.6	6.3
Deaths (people)	13	5	0
Missing (people)	0	0	0
Transfers (10,000 people)	21.3	126.4	4.7
Collapsed houses (10,000)	0.1	0.2	0
Affected area (10,000 hectares)	6.4	23.2	1.23
Direct Economic Losses (Billions of CNY)	27.36	13.29	0.42

The typhoons Hato, Mangkhut, and Higos share similarities: they occurred consecutively, hit Guangdong Province with typhoon-level intensity, and exhibited distinct seasonal patterns. Their similar tracks suggest Guangdong, particularly Zhuhai, is a key impact zone. Despite these similarities, they differed in their impact on Guangdong. Based on statistical data regarding the affected population and areas, typhoon Hato demonstrates an exceptiona **l**y high level of destructiveness and ferocity. Nevertheless, it is noteworthy that the direct economic losses attributed to Hato are merely half of those incurred by Mangkhut, along with a relatively lower count of fatalities. Similarly, despite its slightly reduced intensity, Higos has incurred lesser losses, leading us to inquire about the presence of other influencing factors. These unusual phenomena have piqued our curiosity, necessitating a further exploration of the underlying reasons for such outcomes.

METHODOLOGY

The data collection process involves systematically scraping web content using the Octopus Crawler. Octopus Crawler analyzes web pages and, leveraging regular expressions and XPath techniques, accurately locates and extracts HTML elements containing target data from the web pages. Data pre-processing includes manual deduplication and filtering of invalid content. In the deduplication process, user IDs, posting times, and text contents are compared to effectively remove duplicate Weibo entries. Invalid content filtering removes irrelevant items such as advertisements, entertainment news, and posts containing only emoticons, symbols, and web links with minimal text information, thereby improving the purity and accuracy of the data. Next, the preprocessed dataset is utilized for keyword extraction using the "jieba" package and TF-IDF (Term Frequency-Inverse Document Frequency) algorithm in Python 3.11. The "jieba" package performs Chinese word segmentation, laying the foundation for TF-IDF calculation. TF-IDF takes into account both term frequency (TF) and inverse document frequency (IDF), where TF reflects the importance of a word in the current document (Equation (1)), while IDF reflects its generality in the corpus (Equation (2)). The TF-IDF value is obtained by multiplying TF and IDF (Equation (3)), with a higher value indicating greater importance of the word. Where n represents the frequency of word t in document d, N stands for the total number of words in document d, p denotes the total count of documents in corpus D, and q indicates the number of documents that contain word t.

$$\mathrm{TF}(t, d) = \frac{\mathrm{n}}{\mathrm{N}} \quad (1)$$

$$\mathrm{IDF}(t, D) = log\ \frac{p}{\mathrm{q} + 1} \quad (2)$$

$$\mathrm{TF} - \mathrm{IDF}(\mathrm{t}, \mathrm{d}, \mathrm{D}) = \mathrm{TF}(t, d) \times \mathrm{IDF}(\mathrm{t}, \mathrm{D}) \quad (3)$$

SnowNLP is a package in Python 3.11 for Chinese text sentiment analysis. It assigns a sentiment score between 0 and 1, with values closer to 1 indicating a more positive sentiment, closer to 0 denoting a more negative sentiment, and around 0.5 suggesting neutrality or ambiguity. The score is often discretized into positive, neutral, and negative categories for easier interpretation in practical applications.

RESULTS

Major Events During Typhoon Processes

Utilizing big data technology, a systematic collection and chronological arrangement of key event news is conducted for the 72 hours surrounding typhoon landings, encompassing sources such as Zhuhai meteorological websites. This encompasses timely warnings, government emergency responses, societal participation, mutual assistance, and post-disaster recovery. Through in-depth analysis, a comprehensive understanding of typhoon emergency management is achieved. Specifically, by comparing response effectiveness, practical strategies are identified, and areas for improvement are pinpointed. In responding to the typhoon, Zhuhai has adopted a series of effective and strategic emergency response initiatives (Table 2). Typhoons Hato, Mangkhut and Higos have some similarities in Zhuhai's emergency management initiatives, including the following:

1. Early warning system activation: Before typhoon strikes, Zhuhai city has activated the early warning system, releasing early warning information to the public through radio, television, cell phone text messages and other means to remind the public to take precautionary measures.
2. Emergency plan initiation: Before the typhoon struck, Zhuhai all activated their emergency plans, taking measures such as stopping work, production, market and classes, opening a number of shelters, relocating people from areas likely to be affected, carrying out hidden dangers, and taking precautionary and coping measures to minimize the impact of the typhoon on production and life.
3. Organization of rescue forces: In the wake of typhoons, Zhuhai has promptly organized rescue forces, including firefighting, medical and public security departments, to be ready to respond to emergencies and to actively carry out rescue work to ensure that the affected citizens receive timely relief.
4. Post-disaster recovery work: In the aftermath of typhoon, Zhuhai acted swiftly and actively carried out post-disaster recovery work. Government staff, the public, party members and volunteers from all sectors of society made concerted efforts to work together on infrastructure repair, road clearing, food safety inspections, etc., to ensure that citizens' living and working order returned to normal as soon as possible.

Comparative analysis of typhoons Mangkhut and Higos shows earlier warnings, leading to better Zhuhai preparation. Government, organizations, and public responded swiftly and effectively during critical 24 hours, with district governments actively identifying risks. This improved typhoon response in Zhuhai is not coincidental, but a result of past reflections and improvements. Before super typhoon Hato, Zhuhai lacked similar-sized typhoon references, potentially leading to underestimation of the typhoon's hazards by the public. Challenges prompt Zhuhai to improve its response, significantly enhancing warning accuracy and timeliness. The emergency response becomes more efficient, with departments quickly coordinating resources and protecting infrastructure. For instance, during Higos, a comprehensive emergency command center ensures timely and accurate warning dissemination.

Table 2. Timeline of emergency management measures for typhoon Hato, Mangkhut, Higos

	Hato	Mangkhut	Higos
Pre-72-48h	/	● Disaster warning ① Typhoon White Warning issued	● Disaster warning ① Tropical Depression forcast issued
Pre-48-24h	● Disaster warning ① Typhoon White Warning issued ② Typhoon Orange Warning issued ③ Typhoon Blue Warning issued ● Governmental Emergency response ① Level IV Emergency response activated	● Disaster warning ① Typhoon Blue Warning issued ② Typhoon Yellow Warning issued ● Governmental Emergency response ① Level II Emergency response activated	● Disaster warning ① Typhoon White Warning issued

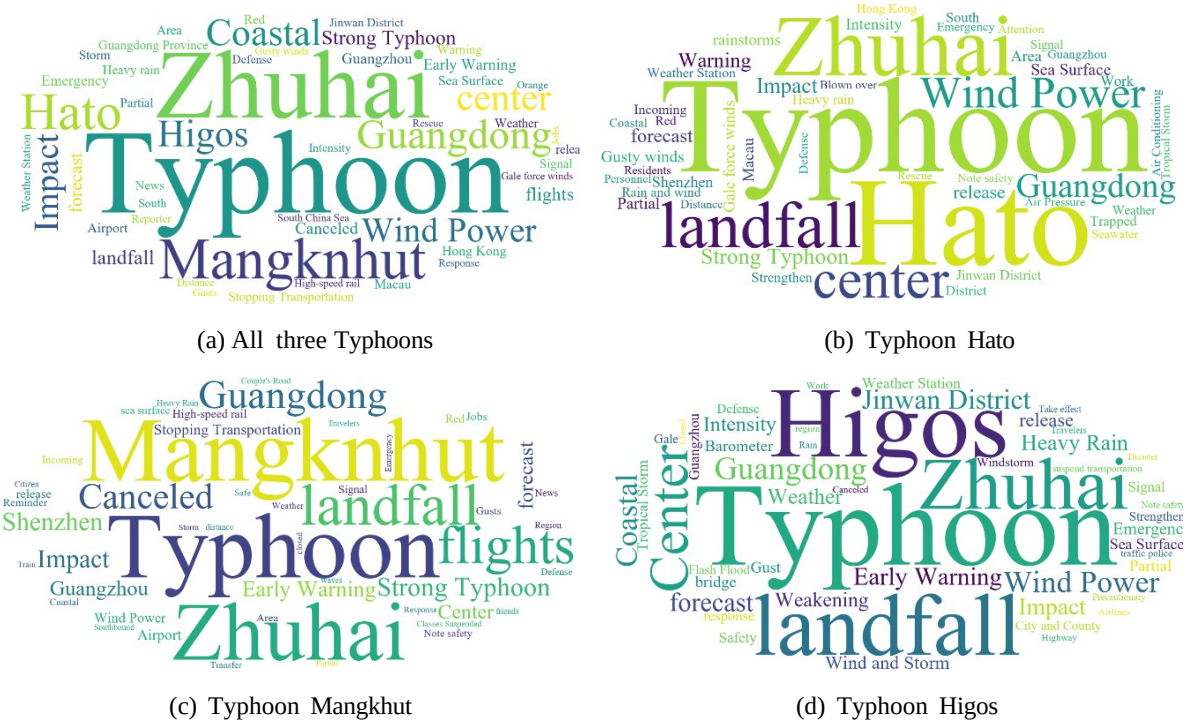
	Hato	Mangkhut	Higos
Pre-24h	● Disaster warning ① Typhoon Yellow Warning issued ② Typhoon Red Warning issued ③ Rainstorm Orange Warning issued ④ Rainstorm Yellow Warning issued ⑤ Wave I Alert (Red) and Storm Surge II Alert (Orange) issued ● Governmental Emergency response ① Level III Emergency response for typhoon prevention activated ② Level I Emergency response for typhoon prevention activated ● Social Emergency response ① Stopping work, business, market and school ② Controlling traffic control successively ③ Transferring people in close proximity to danger and opening shelters ④ Pre-positioning of rescue and relief materials and staff	● Disaster warning ① Typhoon Red Warning issued ② Rainstorm Red Warning issued ● Governmental Emergency response ① Level I Emergency response for typhoon prevention activated ② Level I Emergency response for flood activated ③ "Emergency Mobilization Order to All Citizens on typhoon" issued ④ Level II Emergency response for disaster issued ● Social Emergency response ① Stopping work, business, market and school ② Controlling traffic control timely ③ Transferring people in close proximity to danger and opening shelters ④ Pre-positioning of rescue and relief materials and staff ⑤ Detecting potential danger	● Disaster warning ① Typhoon Blue Warning issued ② Typhoon Yellow Warning issued ③ Typhoon Orange Warning issued ④ Typhoon Red Warning issued ⑤ Rainstorm Yellow Warning and Thunderstorm Warning issued ⑥ Rainstorm Red Warning issued ● Governmental Emergency response ① Level IV Emergency response for typhoon and flood prevention activated ② Level III Emergency response for typhoon prevention activated ③ Level I Emergency response for typhoon prevention activated ● Social Emergency response ① Stopping work, business, market and school ② Controlling traffic control timely ③ Transferring people in close proximity to danger and opening shelters ④ Pre-positioning of rescue and relief materials and staff ⑤ Detecting potential danger ⑥ Government departmental coordination and emergency management system

	Hato	Mangkhut	Higos
24h	● Disaster warning ① Typhoon Blue Warning issued ② Rainstorm Warning cancelled ③ Typhoon Warning cancelled ● Social Emergency response ① Traffic unblocked gradually ② Emergency rescue and repair ③ Road clearing	● Disaster warning ① Typhoon Orange Warning issued ② Typhoon Red Warning issued ● Social Emergency response ① Traffic unblocked gradually ② Emergency rescue and repair ③ Road clearing	● Disaster warning ① Typhoon Orange Warning issued ② Typhoon Blue Warning issued ● Social Emergency response ① Traffic unblocked gradually ② Emergency rescue and repair ③ Road clearing
24-48h	/	● Disaster warning ① Typhoon Warning cancelled	/
48-72h	/	/	/

Word Cloud Analysis

Public attention and participation in typhoon disasters play a vital role in emergency management. Employing data scraping techniques on the Weibo platform, we captured and analyzed public Weibo postings spanning the entire three days before and after each typhoon landfall, inclusive of details such as post timestamps, authors, and content. This yielded a rich trove of raw data for subsequent analysis, enabling an in-depth exploration of public focus points, emotional attitudes, and shifts in social discourse during typhoon events. After meticulous screening and organization, 11,221 Weibo posts were retained from an initial pool of 12,924. Of these, 4,693 pertained to typhoon Hato (2017), 4,862 to typhoon Mangkhut (2018), and 1,666 to typhoon Higos (2020).

Employing the "jieba" package and TF-IDF calculations for keyword extraction, we observed that keywords with higher term frequencies typically exhibit correspondingly elevated TF-IDF values, substantiating the efficacy of our keyword extraction approach. A threshold of 0.1 is chosen, selecting keywords as those with TF-IDF values exceeding this threshold. The Chinese keywords are then manually translated into English and used to generate a word cloud via Python programming. This allowed us to gain a visual, intuitive understanding of the focal points of public attention and emotional responses during the typhoon period. Statistical analysis of Weibo posts on the three typhoons identified high-frequency words (Figure 4a) like "typhoon," "Zhuhai," "landfall," "news," and "warning," reflecting the substantial societal impact and concern for the affected region. Meanwhile, each typhoon showed distinct focal points: typhoon Hato in 2017 (Figure 4b) emphasized intensity, urgency, and severe damages; Mangkhut in 2018 (Figure 4c) highlighted the importance of preparedness, safety measures, and post-disaster rescue; Higos in 2020 (Figure 4d) underscored the role of early warning systems in risk mitigation and loss reduction. These findings demonstrate continuous improvements in typhoon emergency management in terms of technology, response maturity, risk management strategies, social engagement, and public awareness.



(a) All three Typhoons (b) Typhoon Hato

(c) Typhoon Mangkhut (d) Typhoon Higos

Figure 4. Weibo content keywords about typhoons

Weibo Keyword Classification Statistics Analysis

Weibo user content is categorized as Warning (disaster alerts and prevention), Damage (post -landfall situations), and Response (immediate rescue and aid). Figure 5 highlights the stage characteristics in the 144 hours surrounding typhoon landfall. As typhoons approach, content shifts from warnings to damage reports and response efforts. The number of posts peaks post-landfall, reflecting societal concern and urgent response. Before landfall, Weibo disseminates warnings to prepare the public. Afterwards, damage reports and rescue actions dominate, with posts peaking, reflecting societal urgency. As recovery begins, content transitions to reflecting progress, unity, and hope for the future. The process of discussions, ranging from early warnings to disasters and then to rescue efforts, reflects the evolution of typhoon characteristics and the shift of public attention. Furthermore, comparing the three typhoons, Typhoon Mangkhut had the highest information volume, with a significant focus on warnings around landfall. In contrast, Typhoon Hato mainly comprised damage messages, possibly due to lessons learned from the previous year, leading to increased emphasis on pre-disaster warnings and mitigation, which may have contributed to reduced losses during Typhoon Mangkhut. The lower volume and attention for Typhoon Higos may be attributed to its milder impact, swift movement, and the accumulation of typhoon experience, enabling improved early warning and response. Social media is vital in typhoon management, providing real-time information and mobilizing the public.

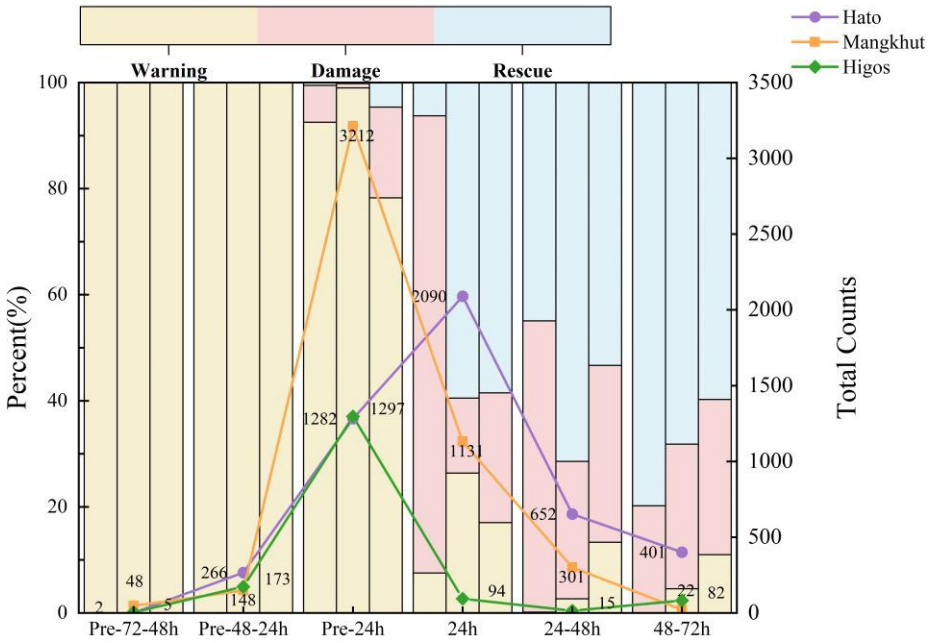


Figure 5. Weibo messages classification and counts

(Left to right: typhoon Hato, Mangkhut, Higos. Bar show percentages, Line are number of Weibo posts)

Public Sentiment Change Analysis

SnowNLP package is used for sentiment analysis, classifying Weibo content as positive, negative, or neutral based on sentiment scores (scores > 0.6 = positive, < 0.4 = negative, 0.4 -0.6 = neutral). Figure 6 reveals how Weibo reflect public emotional responses to approaching typhoons. Prior to a typhoon, positive emotions tend to gradually decrease or remain stagnant, while negative emotions escalate, indicative of rising anxiety and concern. During the typhoon, negative sentiments peak due to fear and distress caused by the storm's impact and hazards. Post-typhoon, positive emotions rebound and negative ones subside as individuals recover and appreciate relief efforts. By comparative analysis, typhoon Higos exhibits a comparatively high level of positive sentiment, paralleling its relatively low disaster intensity and damage, along with more structured and efficient emergency management strategies. Furthermore, there has been a decline in the proportion of negative sentiment and a concurrent increase in positive sentiment from 2017 to 2020, suggesting that the public has become more proficient in typhoon emergency response. This accumulated experience not only assists individuals in overcoming their fears but also testifies the continuous improvement in emergency mana gement practices.

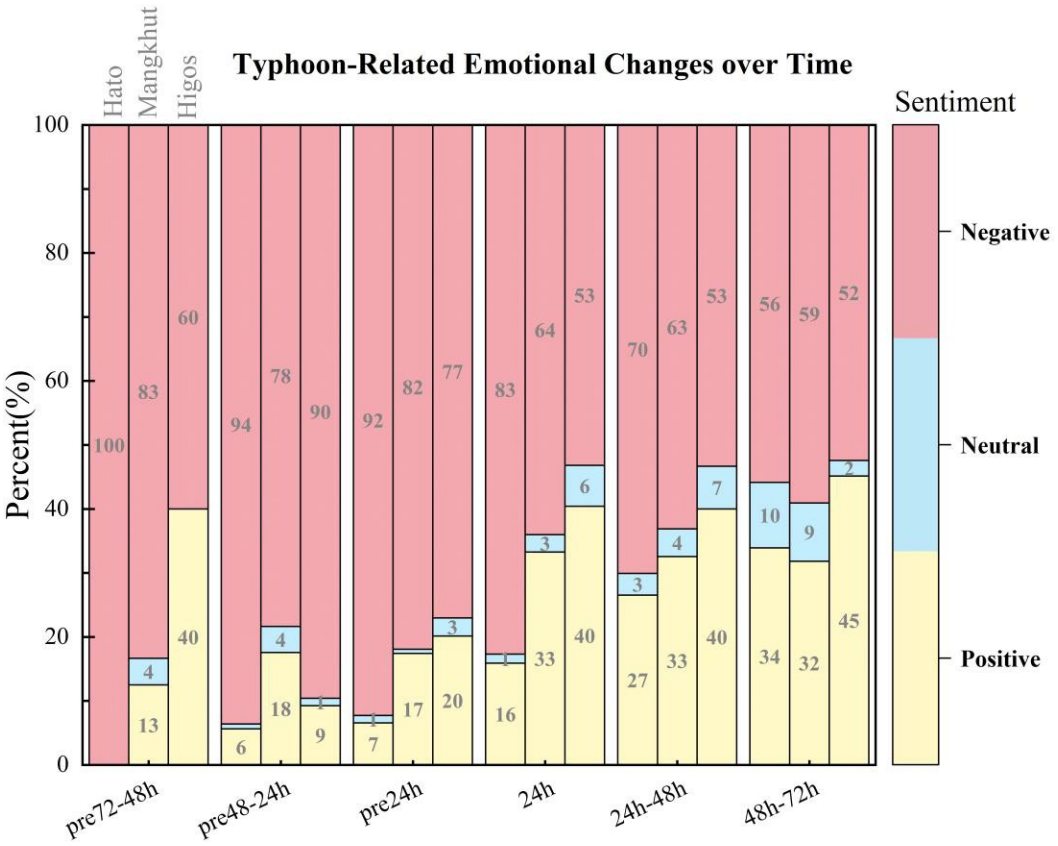


Figure 6. Sentiment evolution in public Weibo activity during typhoons

(Left to right: typhoon Hato, Mangkhut, Higos)

DISCUSSION

This study presents a comprehensive examination of the impacts of typhoons Hato (2017), Mangkhut (2018), and Higos (2020) on Zhuhai, incorporating an analysis of typhoon characteristics, response strategies, and public sentiment. Initially, comparative analysis of typhoon paths and collation of key disaster data facilitated a holistic understanding of the typhoons and their effects on Zhuhai. Subsequent scrutiny of implemented response measures post-landfall revealed a decreasing trend in economic losses and casualties, indicating enhanced emergency management capabilities and accumulated experience. Weibo data analysis was employed to gauge public sentiment and concerns. The keyword clouds generated from this analysis reveal distinct foci of public attention, ranging from the intensity of Hato, to the rescue efforts during Mangkhut, and the warning sy stems associated with Higos. Statistical analysis of content categories elucidated the shifting public attention, encompassing warnings, disaster impacts, and rescue operations. Sentiment analysis of Weibo showed negative emotions prevailing during typhoons, gradually giving way to positive sentiments as focus shifted towards recovery. A declining trend in negative emotions from 2017 to 2020 suggests improving emergency management practices.

All typhoons have provided valuable experiences and lessons to Zhuhai, leading to enhancements in typhoon prevention mechanism. The Hato specifically highlighted the seriousness and complexity of typhoon disasters, resulting in significant improvements. Zhuhai has upgraded its early warning system, forecast accuracy, and emergency response, reflected in the "Zhuhai Typhoon Prevention Regulations" since February 2020. Technological advancements have strengthened Zhuhai's emergency management. Media and edu cation have raised public awareness, fostering community-wide disaster prevention and government cooperation. Despite challenges from climate change and urbanization, Zhuhai's progress highlights the need for enhanced early warning systems, cross-sectoral coordination, infrastructure, and social participation in disaster response. Zhuhai's achievements in typhoon prevention underscore the significance of experience summarization, mechanism building, and public awareness in tackling natural disasters. Sharing effective emergency management practices globally is crucial to reducing typhoon threats.

LIMITATIONS

When conducting a thorough examination of the study, several limitations must be acknowledged, which may potentially impact the accuracy and comprehensiveness of our findings. Firstly, technological and platform constraints limit data collection to hourly intervals and a maximum of 50 pages, possibly omitting crucial content and affecting the comprehensiveness and depth of the data. Secondly, the classification of Weibo content involves subjective judgments, such as simplifying complex content related to disasters and rescue into a single category, which may affect the accuracy of the results. Moreover, In employing SnowNLP for sentiment analysis, its limitations primarily stem from inadequate handling of emotional complexity. Social media texts often exhibit user sentiments that are not simply dichotomized as positive, neutral, or negative, but rather may encompass various compounded emotions interwoven within a single post. General-purpose sentiment models like SnowNLP may struggle to discern subtle emotional variations in public responses to typhoon events. Categorical sentiment classification (positive, neutral, negative) as opposed to fine-grained scoring, sacrifices analytical precision. Despite these limitations, this study still provides meaningful analysis and conclusions based on the available data.

CONCLUSIONS

Typhoons, characterized by immense destructiveness and unpredictability, pose substantial threats to human life, infrastructure, and societal stability. While the extent of damage varies for typhoons of comparable intensity and trajectory, efficacious emergency management is vital. Essential components include early warnings, evacuations, preparedness training, and swift post-event recovery, which collectively mitigate impacts, minimize casualties, and enhance community resilience.

This study examines the coastal city of Zhuhai, China, frequently battered by typhoons, comparing its emergency management during Hato (2017), Mangkhut (2018), and Higos (2020). Employing a big data approach, it collates news, Weibo posts, and ancillary data preceding and following each typhoon's landfall. Key event sequences for 72 hours around landfall and word frequency analysis, categorical statistics, and sentiment analysis on Weibo contents are conducted to trace the emergency management process, analyze Zhuhai's measures, characteristics, and effectiveness. Zhuhai demonstrates continuous adaptation and optimization of its response to similarly intense typhoons. The study draws the following main conclusions: Enhanced warning systems significantly boost public responsiveness, curbing disaster losses. Systematic coordination in emergency management is pivotal for seamless rescue and recovery. Heightened public awareness and social engagement are instrumental in loss mitigation.

The study illuminates Zhuhai's strategic adaptations in typhoon response, offering a critical lens on its coping mechanisms. These experiences and strategies provide instructive references for Zhuhai and other coastal cities to better prepare for and mitigate typhoon impacts. The derived lessons hold global significance in strengthening coastal cities' disaster response capabilities, reducing typhoon-related losses, protecting lives and assets, and fostering sustainable urban development.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (Grant No. 72304038), Guangdong Basic and Applied Basic Research Foundation (Grant No. 2022A1515110049), Young Innovative Talents Project from Department of Education of Guangdong Province (Grant No. 2022KQNCX157).

REFERENCES

Anping, P. (2010). The applications of maximal covering model in Typhoon Emergency shelter Location Problem. IEEE.

Athanasia, N., & Stavros, P. T. (2015). Twitter as an instrument for crisis response: The Typhoon Haiyan case study.

Chan, E., Man, A., Lam, H., Chan, G., Hall, B., & Hung, K. (2019). Is Urban Household Emergency Preparedness Associated with Short-Term Impact Reduction after a Super Typhoon in Subtropical City? *International Journal of Environmental Research and Public Health* , *16*(4).

Chen, Q. J., He, Y. R., He, T. T., & Fu, W. J. (2020). The typhoon disaster analysis emergency response system based on uan remote sensing technology.

Choy, C. W., Insurers, T. H. K. F. O., Wu, M. C., & Lee, T. C. (2020). Assessment of the damages and direct economic loss in Hong Kong due to Super Typhoon Mangkhut in 2018.

Ermino, R. P., Abacan, R. C. A., Diamante, N. G. C., Faca, K. M. N., & Juntereal, T. E. D. (2022). Evaluating Urgency of Typhoon-Related Tweets Through Sentiment Analysis Using Artificial Neural Networks. *International Journal of Computing Sciences Research* , 940-950%V 946. //stepacademic.net/ijcsr/article/view/289

G.C, D., Ladas, A., Sambo, Y. A., Pervaiz, H., Politis, C., & Imran, M. A. (2019). An Overview of Post-Disaster Emergency Communication Systems in the Future Networks. *IEEE Wireless Communications*, *26*(6), 132-139. https://doi.org/10.1109/MWC.2019.1800467

Geiger, T., Guetschow, J., Bresch, D. N., Emanuel, K., & Frieler, K. (2021). Double benefit of limiting global warming for tropical cyclone exposure. *Nature Climate Change*(10), 11.

Gettelman, A., Bresch, D. N., Chen, C. C., Truesdale, J. E., & Bacmeister, J. T. (2018). Projections of future tropical cyclone damage with a high-resolution global climate model. *Climatic Change*.

Guidotti, T. L. (2006). Hurricane Katrina: an American tragedy. *Occup Med*, *56*(4), 222-224.

Huo, Z., Duan, W., & Zhou, F. (2019). Ensemble Forecasts of Tropical Cyclone Track with Orthogonal Conditional Nonlinear Optimal Perturbations. *Advances in Atmospheric Sciences*, *36*(2), 231-247.

Jiang, G. Q., Xu, J., & Wei, J. (2018). A Deep Learning Algorithm of Neural Network for the Parameterization of Typhoon‐Ocean Feedback in Typhoon Forecast Models. *Geophysical Research Letters*.

Lai, T. I., Shih, F. Y., Chiang, W. C., Shen, S. T., & Chen, W. J. (2003). Strategies of Disaster Response in the Health Care System for Tropical Cyclones: Experience Following Typhoon Nari in Taipei City. *10*(10), 1109-1112.

Li, M., Wu, W., Wang, J., Che, Z., & Xie, Y. (2016). Simulating and mapping the risk of surge floods in multiple typhoon scenarios: a case study of Yuhuan County, Zhejiang Province, China. *Stochastic Environmental Research and Risk Assessment*, 1-15.

Li, Y. A. W., Jidong %A Tang, Rumei %A Wu, Kejie %A Nie, Juan %A Shi, Peijun %A Li, Ning %A Liu, Lianyou. (2022). Vulnerability to typhoons: A comparison of consequence and driving factors between Typhoon Hato (2017) and Typhoon Mangkhut (2018). *Science of The Total Environment*, *838*, 156476. https://doi.org/10.1016/j.scitotenv.2022.156476

Liu, T., Zhang, H., & Zhang, H. (2020). The Impact of Social Media on Risk Communication of Disasters—A Comparative Study Based on Sina Weibo Blogs Related to Tianjin Explosion and Typhoon Pigeon. *International Journal of Environmental Research and Public Health* , *17*(3), 883. https://www.mdpi.com/1660-4601/17/3/883

Liu, Z.-g., Li, X.-y., Durrani, D. K., & Zhang, R.-j. (2023). Case-driven assessment of power grid emergency response capability considering typhoon disaster scenario diversity [Article]. *Journal of Cleaner Production*, *386*, Article 135813. https://doi.org/10.1016/j.jclepro.2022.135813

Lu, X., Yu, H., Ying, M., Zhao, B., Zhang, S., Lin, L., Bai, L., & Wan, R. (2021). Western North Pacific Tropical Cyclone Database Created by the China Meteorological Administration. *Advances in Atmospheric Sciences*, *38*(4), 690-699. https://doi.org/10.1007/s00376-020-0211-7

Pan, W., Zhong, S., Huang, S., & Wu, D. (2017). Online Power Grid Emergency Material Demand Forecasting Model in Typhoon Disaster Based on Relevance Vector Machine. International Symposium on Computational Intelligence & Design.

Pollack, J., Thompson, C., Tran, M., Graham, C., & Wolcott, M. (2015). A guide to social media emergency management analytics: Understanding its place through Typhoon Haiyan tweets. *Statistical Journal of the IAOS: Journal of the International Association for Official Statistics* .

Qin, L., Xu, W., Zhao, X., & Ma, Y. (2019). Typhoon track change-based emergency shelter location-allocation model: a case study of Wenchang in Hainan province, China. *Injury prevention*, *ePub*(ePub), ePub-ePub. http://dx.doi.org/10.1136/injuryprev-2018-043081

Santiago, J. S. S., Manuela, W. S., Tan, M. L. L., Sa?Ez, S. K., & Tong, A. Z. U. (2016). Of timelines and

timeliness: lessons from Typhoon Haiyan in early disaster response. *Disasters*, n/a-n/a.

Santos, C. T., Toda, L., Orduña, J. R., Santos, F. D., & Ferrão, J. (2016). The impacts of Typhoon Haiyan in the Philippines: Implications to land use planning.

Shiyuan, X. U., Jun, W., Chun, S., Jianping, Y. (2006). Research of the Natural Disaster Risk on Coastal Cities *61*(2), 127-138.

Su, M., Wu, J., & Feng, M. (2022). Responses of Urban Ecosystem Vulnerability and Restoration Assessment on Typhoon Disaster: A Case Study of Zhuhai City, China. *Journal of Sensors*, *2022*, 8344619. https://doi.org/10.1155/2022/8344619

Su, S. S. (2013). The Effect of Tropical Cyclones (Typhoons) on Emergency Department Visits. *The Journal of Emergency Medicine*.

Sun, D., Zhang, L., & Su, Z. (2020). Evacuate or Stay? A Typhoon Evacuation Decision Model in China Based on the Evolutionary Game Theory in Complex Networks. *International Journal of Environmental Research and Public Health*, *17*(3).

Sylwanowicz, L., Schreiber, M., Anderson, C., Gundran, C. P. D., & Tuazon, A. C. (2018). Rapid Triage of Mental Health Risk in Emergency Medical Workers: Findings From Typhoon Haiyan – CORRIGENDUM. *Disaster Medicine and Public Health Preparedness*, *12*(6), 1.

Uichanco, J. (2022). A Model for Prepositioning Emergency Relief Items Before a Typhoon with an Uncertain Trajectory. *Manufacturing and service operations management* (2), 24.

Wang, K., Yang, Y., Reniers, G., & Huang, Q. (2021). A study into the spatiotemporal distribution of typhoon storm surge disasters in China. *Natural Hazards*, 1-20.

Wang, X., Choi, T.-M., Liu, H., & Yue, X. (2018). A Novel Hybrid Ant Colony Optimization Algorithm for Emergency Transportation Problems During Post-Disaster Scenarios. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *48*(4), 545-556. https://doi.org/10.1109/TSMC.2016.2606440

Wei, Chen, Xiaozhong, Zhang, Zhengxie, Zhou, Jianrong, He, Hui, & Tao. (2018). Emergency management of reservoirs and water treatment plants in typhoon season:a case study of Ningbo. *Chinese Journal of Population,Resources and Environment, v.16*(04), 80-89.

Xu, H., Lin, N., Huang, M., & Lou, W. (2020). Design Tropical Cyclone Wind Speed when Considering Climate Change. *Journal of structural engineering*, *146*(5), 04020063.04020061-04020063.04020013.

Yin, J., Yin, Z., & Xu, S. (2013). Composite risk assessment of typhoon-induced disaster for China's coastal area. *Natural Hazards*(69-3).

Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., Zhu, Y., & Chen, D. (2014). An Overview of the China Meteorological Administration Tropical Cyclone Database. *Journal of Atmospheric and Oceanic Technology*, *31*(2), 287-301. https://doi.org/https://doi.org/10.1175/JTECH-D-12-00119.1

Zhang, H., Liu, X., Wu, R., Liu, F., & Zhang, W. (2019). Ocean Response to Successive Typhoons Sarika and Haima (2016) Based on Data Acquired via Multiple Satellites and Moored Array. *Remote Sensing*, *11*(20), 2360.

Forced Displacement Situation Awareness with Large Language Models: Ukrainian Refugees in Poland Case Study

Brian Tomaszewski
Rochester Institute of Technology
bmtski@rit.edu

Nataliya Shakhovska
Lviv Polytechnic National University
nataliya.b.shakhovska@lpnu.ua

Harley Emery
University of Maryland
harleyem@umd.edu

Paweł Śniatała
Poznań University of Technology
pawel.sniatala@put.poznan.pl

ABSTRACT

We evaluated the utility of the large language model (LLM) ChatGPT to develop situation awareness related to the forced displacement of Ukrainian refugees into Poland. Utilizing text messages derived from the Help for Ukrainians in Poland Telegram message group, we used ChatGPT to translate messages in multiple languages into English and identify message topics and themes. Topics and themes from beginning of the war in Ukraine in 2022 were analyzed and visualized using K-means clustering and word clouds. The language identification and translation capabilities of the LLM were evaluated by two human evaluators and measured using Kappa (0.86) and BLUE scores (0.46) with the LLM performing effectively. We conclude that the ability of LLM using carefully developed language prompts for large data volume analysis with no need for manual human analysis shows promise for humanitarian analytics focused on rapidly identifying potential key trends, needs, and locations of displaced people.

Keywords

Ukrainian refugees, Poland, Large Language Model, ChatGPT.

INTRODUCTION

The unprovoked Russian invasion of Ukraine in 2022 has created one of the largest humanitarian crises in modern times. According to the UN, as of August 2022, 1,388,339 Ukrainian refugees have registered for temporary protection in Poland (United Nations High Commissioner for Refugees (UNHCR), 2022b) (Figure 1).

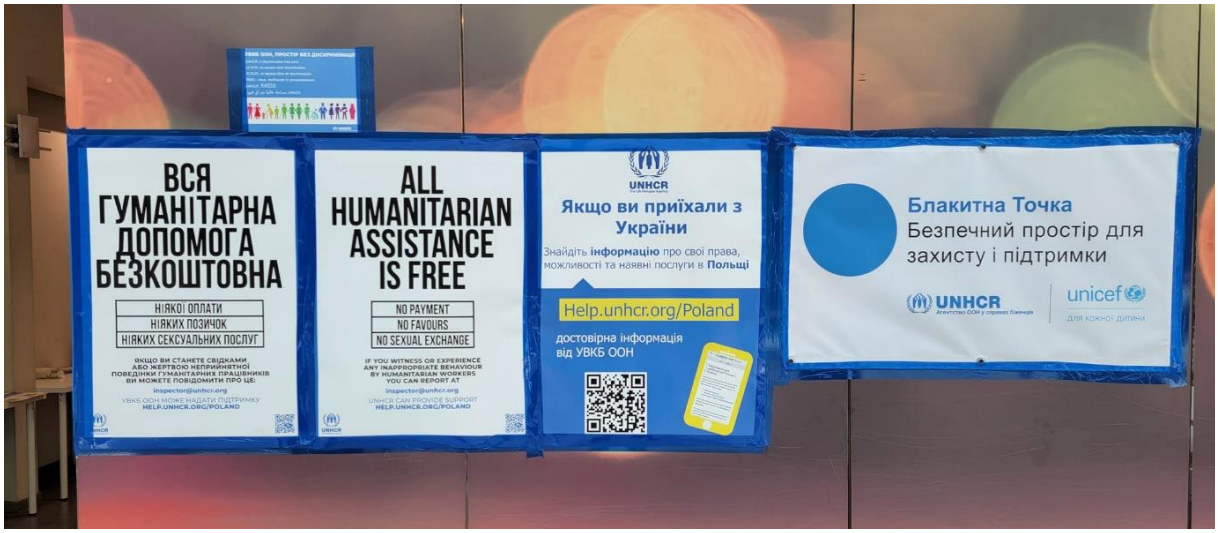


Figure 1. An example of 'Blue Point' refugee assistance center in Poznań, Poland (picture by author, 2022).

Although many Ukrainian refugees are returning home, with the war in Ukraine showing no end in sight, it is likely that many will remain in Poland for many years to come, as is typical in protracted forced displacement situations (UNHCR Regional Bureau for Europe Data Identity Management and Analysis Unit (DIMA), 2022).

International organizations that track forced displacement situations like the International Organization for Migration (IOM, UN Migration) monitor the situation of refugees once they arrive in a host country (c.f (International Organization for Migration, 2023b). In the case of the Ukrainian situation in Poland, it is important to know where and why refugees are settling in particular places in Poland and issues they are facing as there are several issues that may drive these decisions. For example, International Organization for Migration (2023a) outlines how in Warsaw, the main factors for why refugee population chose to settle in Warsaw include employment opportunities, accommodation availability and cost, and availability of health, educational, and other services.

Additionally, the general relationship between Polish society and Ukrainian refugees in Poland overtime is important to monitor. Poland received great praise from the international community for its support of Ukrainians in the first few months of conflict in 2022 in terms of social support, legal status, education and other factors (United Nations High Commissioner for Refugees (UNHCR), 2022a). However, as the war in Ukraine continues into the foreseeable future, some sources are signaling that Polish support systems have already, and will continue, to come under strain (Reidy, 2022).

A significant challenge exists for keeping track of the situations that individual displaced people may be experiencing once inside a host country. For example, official Polish government data sources on PESEL registrations (a system to track foreigners in Poland, see: (gov.pl, n.d.)) can give an indication of where refugees may be settling at the administrative level 2 in Poland (also known as a Powiat, similar to a county or district). Although this is a very useful starting point, refugees do not always reside in the location where they register for PESEL, making this data source not necessarily an accurate representation of what displaced people may be experiencing (Figure 2).

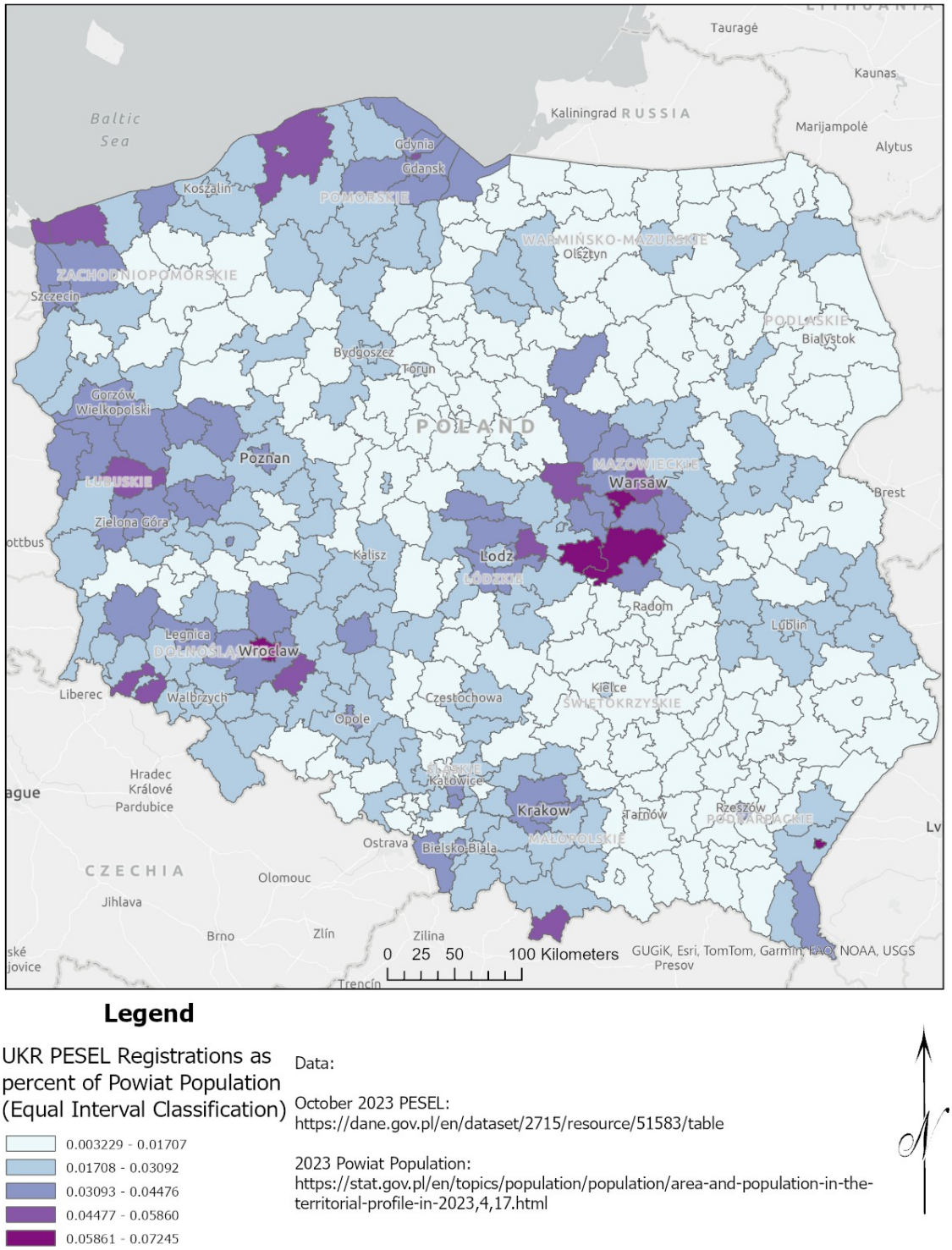


Figure 2. Ukrainian PESEL registrations in Poland as of October 2022. Understanding of where Ukrainian Refugee populations in Poland are is a challenging problem due to data granularity and temporality issues. A map like this is often the best representation of refugee locations but often inaccurate as refugees frequently move after receiving a PESEL.

Besides knowing *where* placed people may be located, are the additional challenges of understanding *what* the issues are that displaced people may be experiencing. Social media has a long established precedent in ISCRAM-related research as one potential source of data that can be used for developing situation awareness about a given crisis (c.f. (Hughes & Palen, 2009; Power, Robinson, & Cameron, 2022).

Generative AI is offering new promise for conducting social media analysis due to the lower entry barriers of accessing Large Language AI models (LLM) using human dialogue prompts. We argue that this approach presents a significant opportunity for non-technical specialists, such as humanitarian professionals, to utilize the power of LLM for social media analytics to ideally fill gaps in situation awareness knowledge about displaced people. A comparative study has been made among Pa LM, BERT and GPT language models for topic categorization using AG News Dataset (Yerramreddy, Marasani, Gowtham, & Abhishek, 2023). However, the investigation for social analysis in Ukrainian language is not given. Thus, in this pilot work we investigated the question: What can LLM reveal about the Ukrainian forced displacement situation in Poland? To investigate this question, we conducted a pilot investigation to specifically investigate and evaluate the use of LLM tools to (1) translate social media messages to English and (2) classify translated social media messages by topic and theme to understand the Ukrainian forced displacement situation in Poland.

BACKGROUND

LLM like ChatGPT are perhaps one of the most significant technology paradigm shifts in in recent years. The ability of the LLM to receive input via human dialogue and provide natural language responses is quickly demonstrating its ability in application domains ranging from business to education. In the humanitarian domain, AI language models can help address several challenges related to natural language processing (NLP)-centered analytics such as language translation and thematic analysis. NLP is generally unknown to humanita rians (Rocca, Tamagnone, Fekih, Contla, & Rekabsaz, 2023). Limited academic research has been conducted to date on the efficacy of LLM models for humanitarian applications. The literature that can be found on the use of AI models like ChatGPT for humanitarian applications has highlighted operational issues such as content creation for advocacy, supporting funding applications, legal document generation, and identifying support organizations (Crew et al., 2023; Roane, 2023). Tanner (2023) identified several issues identified from several humanitarian AI projects related to the broader application of artificial intelligence in humanitarian applications such as effective data strategies, the need for skilled data science professionals and ensuring proper data custody. However, the report explicitly stated that all the projects were conducted before the advent of LLM like ChatGPT.

METHODOLOGY

Figure 3 graphically outlines the workflow used to investigate what LLM can reveal about the Ukrainian forced displacement situation in Poland using social media.

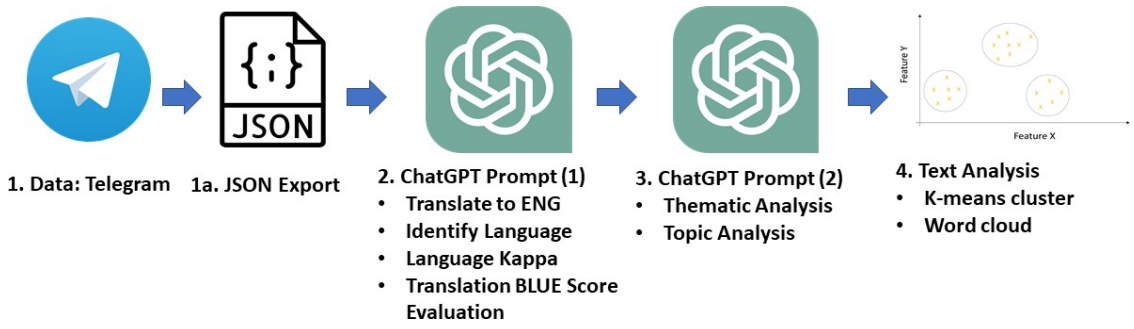


Figure 3. Ukrainian Refugee Social Media Analysis with ChatGPT work flow steps.

Step 1: Data Source

Telegram is a social media platform that is popular in Ukraine. In particular, the service has been extensively utilized by Ukrainians to discuss the ongoing war with Russia and as a tool for information-sharing among Ukrainian refugees in Poland. In the research presented here, we utilized the Telegram message community called Help for Ukrainians in Poland (UKR: Помощь украинцам в Польше) (Telegram, 2024). At the time when this research was conducted, this group had over 17,000 members with frequent daily posting (Figure 4).

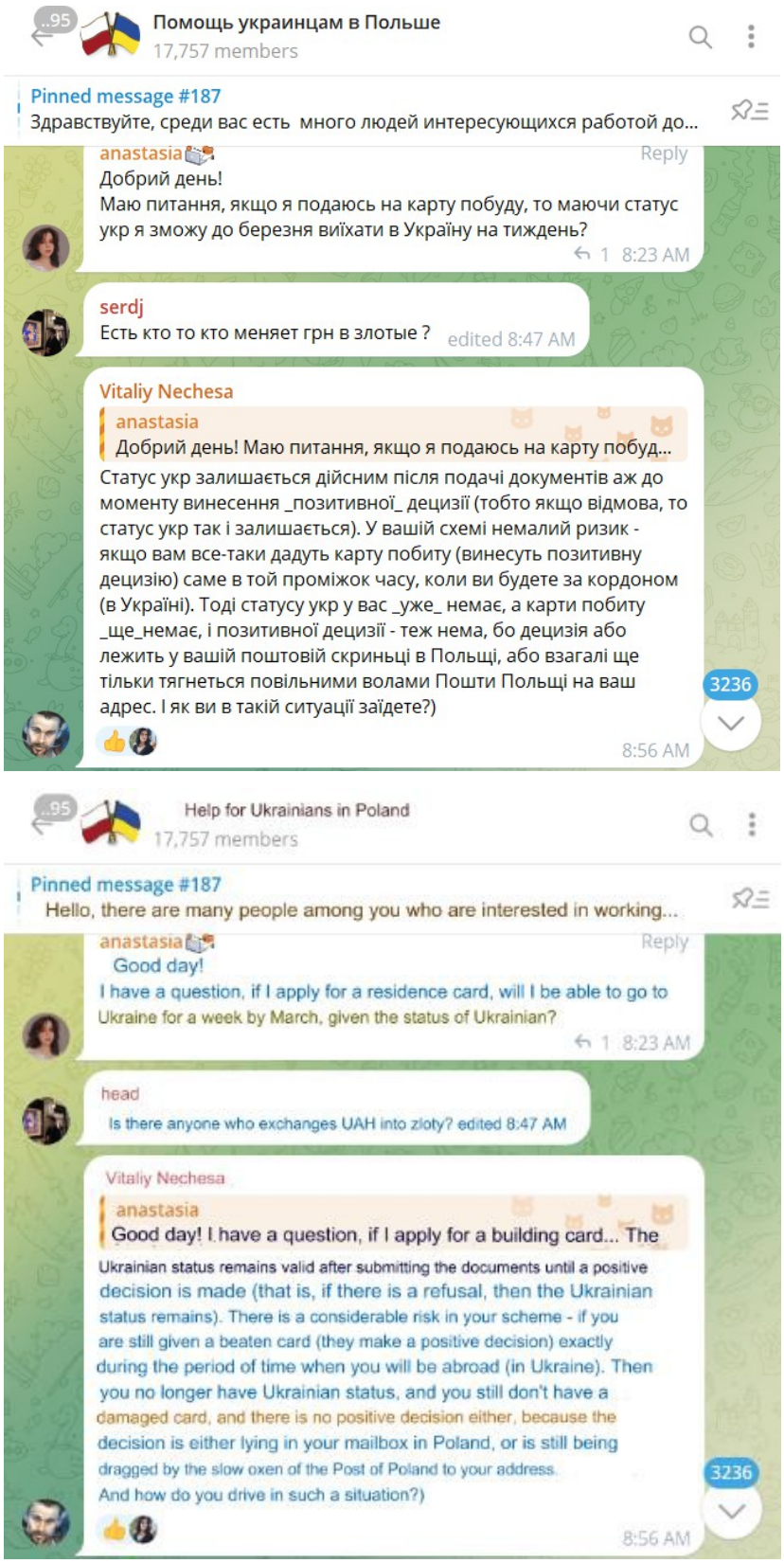


Figure 4. Screen shot of the Help for Ukrainians in Poland Telegram group. The image on the top is in the original source language, the image on the bottom is a Google image translation to illustrate the types of conversions conducted in the group.

Data from Telegram groups can be easily downloaded without requiring any special API programming like other social media platforms Thus, we downloaded the entire collection data from this group dating back to 2022 at the beginning of the war up until January 2024 . This yielded over 2.39 gb of data ranging from images to audio clips to text messages. In the work presented here, we focused only analysis of text messages that were provided in a JSON format as per step 1a seen in Figure 3.

Step 2: ChatGPT Prompt 1 – Message Translations and Language Identification

We utilized the gpt-3.5-turbo-0613 model via the OpenAI API in Python (OpenAPI, 2024). As this was an exploratory pilot project, this model was deemed sufficient vs other alternatives.

Perhaps the most important area of research with LLMs is prompt engineering, or how a human makes a natural language request of the AI model. Engineering plays a crucial role in designing AI systems that can understand and fulfill user queries accurately. By crafting prompts that are specific, well-structured, and tailored to the task at hand, engineers can guide the AI to produce results that meet user expectations. Beyond relevance, engineering also involves formatting the output of AI systems in ways tha t facilitate further analysis and processing. This includes organizing the information, presenting it in a structured manner, and ensuring compatibility with downstream tools and workflows. Careful prompt engineering is essential to ensure that the AI retu rns back relevant results as well as results that are formatted in ways that can be utilized for further analysis. By structuring prompts and results retuned in JSON format, they can be easily manipulated and managed programmatically, allowing for automation and integration with other workflows. This can be particularly useful in scenarios where a large number of prompts must be generated or dynamic adjustment of prompt parameters.

In step 2 of the workflow, we were interested in (1) having the Telegram messages translated to English and (2) have the LLM identify what language it thought the message was in. Our motivation was twofold. First, English is generally the language used by the international humanitarian community, even though there are numerous humanitarian staff in Poland that speak Polish and Ukrainian. Therefore, we wanted to investigate how using LLM could enable a wider group of humanitarian professionals to access information in languages other than English. Second, people from Ukraine in many cases speak both Ukrainian and Russian, or primarily Russian. The messages provided in the Help for Ukrainians in Poland Telegram group board are posted in both a mix of Russian, Ukrainian and other languages. Therefore, we are interested to assess how well the language model was at identifying a particular language in order to build trust that results from the LLM were accurate.

The following prompts were used to receive both the English translation of an input text and the language that ChatGPT identified from its analysis:

PROMPT_PART_1 = "Translate this text to English <start> "

PROMPT_PART_2 = " <end>. Return back two things. The first is your translation to English of text that was between the <start> and <end> tags. The second is a one-word description of the language of text that was between the <start> and <end> tags. "

PROMPT_PART_3 = "Put the two items you return into a JSON structure. Your translation to English of text that was between the <start> and <end> tags placed inside a JSON tag named translation. Your one-word description of the language of text that was between the <start> and <end> tags inside a JSON tag named language. Do not return any additional text, descriptions of your process or information beyond two items and output format of the tags specified."

Telegram message texts was processed from the JSON file downloaded in Step 1A and placed inside of the start and end tags as per the aforementioned prompts. The use of <start> and <end> tags was used after several iterations of developing prompts as they were found to be effective for clearly defining what text chat GTP should be focused on for translations as opposed to output formatting instructions.

LLM models can also benefit from being given a motivation to help establish the task being asked to do. Therefore, the following motivation message was sent in along with the prompts.

MOTIVATION_MESSAGE = "You are a skilled linguist who is an expert in identifying what language a given text is and doing translations of that text to English."

Using these prompts, we processed over 20,000 messages into English dating from 2022 to 2023. All of the English translations along with the original Source message, date of the message ID of the message and identified language of the message were stored in a CSV file. This allowed for the next step in the workflow to conduct thematic and topic analysis on the English translations that we discuss in the next section.

Step 3: ChatGPT Prompt 2 – Thematic and Topic Analysis

Identifying themes and topic from text is a common natural language processing goal particularly in the application of developing situation awareness in an ongoing protracted crisis situation like Ukrainian refugees in Poland (Rocca et al., 2023). Therefore, in step 3 of our process, we investigated the ability of ChatGPT to identify topics and themes from the English language translations that were derived in step 2 and discussed previously. Our motivation here was to investigate how high level thematic and topic summaries of what was being discussed in social media could be derived for developing an initial situation assessment that could then be followed up upon for further deeper investigation and situation awareness at local levels as per the well-established ideas of the process of achieving situation awareness (Endsley, 1995).

The following prompts were used to receive to receive topic and thematic summaries of individual messages:

PROMPT_PART_1 = "Conduct a thematic analysis of this text <start> "

PROMPT_PART_2 = " <end>. Return back two things."

PROMPT_PART_3 = "The first is your general thematic assessment of text that was between the <start> and <end> tags. The second are any topics you found in the text that was between the <start> and <end> tags. "

PROMPT_PART_4 = "Put the two items you return into a JSON structure. Your general thematic assessment of text that was between the <start> and <end> tags placed inside a JSON tag named theme. Any topics you found in the text that was between the <start> and <end> tags inside a JSON tag named topic."

PROMPT_PART_5 = "Do not return any additional text, descriptions of your process or information beyond two items and output format of the tags specified."

Like step 2, English language message texts processed in Step 2 were placed inside of the start and end tags of the aforementioned prompts.

The following motivation was used to help establish the task that ChatGPT was being asked to do.

MOTIVATION_MESSAGE = "You are a skilled humanitarian analyst who is an expert in conducting thematic analysis of English language texts."

Topics and themes returned back from chatGPT as text were stored in a .csv file as an input to topics and theme text analysis in step 4 in the process that we discussed next

Step 4: Topics and Themes Text Analysis

Given the very large volume of unstructured text derived from step 3 thematic and topic analysis, we use the well-established methodology of K-means clustering to group similar topics and themes based on words in those messages to derive a summary representation of topics and themes (Kumbhar, Mhamane, Patil, Patil, & Kale, 2020). Specifically, topic and theme texts were first processed to remove stop words, or words with no significance like ‘the’ ‘and’ ‘is’, and converted into a TD-IDF vector, or the importance of a given word in a single Telegram message relative to the overall collection of messages. The optimal number of clusters were calculated using a silhouette score that determines how well the messages might cluster in general and the optimal number of clusters. K-means clustering, which will partition text found in the messages into similar and dissimilar clusters, was then performed using the optimal number of clusters derived from the silhouette score. The dimensionality of the clusters was then reduced using principal component analysis (PCA). PCA was performed to reduce the dimensionality of the data space (or reducing the number of words in a given cluster while preserving the most important words) to allow for graphical representation of clusters derived from K means processing. A graphical representation of the clusters derived by calculating the centroid of a given cluster based on the term with the highest TF-IDF score in that cluster. This term was then plotted to create a graphical representation of the cluster. Additionally, word cloud generations were generated as an additional source of graphical representation of the text analysis.

RESULTS

As discussed previously, a particular motivation of this work was to utilize LLM to conduct language translations so that content derived from social media could be more widely accessible to a wider range of humanitarian professionals who may not in fact speak Polish, Ukrainian, or Russian. Therefore, we were interested in the effectiveness of the LLM to (1) specifically identify what language a given message was in and (2) the accuracy of the translation from a given language into English. Understanding the accuracy and effectiveness of such language tasks is important so that the AI can be considered trustworthy.

To facilitate these evaluations, two native speakers of Ukrainian and Russian with fluency of English were

recruited to perform the same tasks that the LLM conducted in terms of (1) identifying what language a given message was in and (2) do a translation of the message into English. These two evaluators worked on a sample of 108 messages. This sample size was deemed sufficient given the high expense of conducting human translations. Messages were selected for human evaluation with the criteria that the message contained no emojis or other non-text characters that could not be directly translated into English. Comparison of the human vs LLM of langua ge identification identified using Cohen's Kappa and accuracy of the LLM translation to English evaluated using BLEU (Bilingual Evaluation Understudy) score are discussed next.

Language Identification Kappa

Cohen Kappa is a well-established statistic to measure to assess the level of agreement between two evaluators (Viera & Garrett, 2005). In the work presented here, we were interested in comparing what language a given Telegram message was as rated by both a human evaluator and the LLM. The percentage of agreement between the human annotator and the LLM was 96.2% with a Kappa score of 0.86. These results indicate that the LLM was very effective at identifying what language a given Telegram message was across Russian, Ukrainian, Polish and other languages when compared to the same assessment done by a human.

Translation BLUE Score

BLEU is a very well established metric for automated evaluation of machine translations (Papineni, Roukos, Ward, & Zhu, 2002). English translations of a given Telegram message provided by the LLM were evaluated against two human translations of the same message using a Corpus BLUE score. A 4-gram precision was calculated with the weights 0.1, 0.1, 0.4, 0.4 to place higher emphasis on 3-gram and 4-gram precision (40% each) as these would be considered more useful for thematic and topic analysis and lower on 1-gram and 2-gram precision (10% each). The Corpus BLUE score result was: 0.46. Thus, it can be considered the LLM translations provided "High quality translations" (Google Cloud, n.d.).

K-MEANS TERM CLUSTERING AND WORD CLOUDS

Given the overall time frame of the data set that was processed, it was deemed more feasible to examine thematic and topic analysis based on particular time periods during the forced displacement of Ukrainians into Poland. Specifically, we performed K-means clustering on messages from the date range of 26 – 28 February 2022 within the overall data set processed by the LLM. This date range covers the very beginning of the Help for Ukrainians in Poland telegram group. It also coincides with the opening of the war in Ukraine including the first significant military attacks on large Ukrainian cities such as Kiev and Kharkiv causing thousands of people to begin fleeing to Poland (BBC News, 2022). Our goal was to do an initial assessment of themes that would be found in the data set. Figure 5a and 5b graphically show the results of K-means clustering around thematic analysis conducted by the LLM. Figure 6 shows a word cloud representation of thematic analysis conducted by the LLM. Figure 7 shows the results of a K-means clustering of the topic analysis conducted by the LLM. Figure 8 shows the results of a word cloud representation of the topic analysis conducted by the LLM. These figures are followed by a discussion of these results.

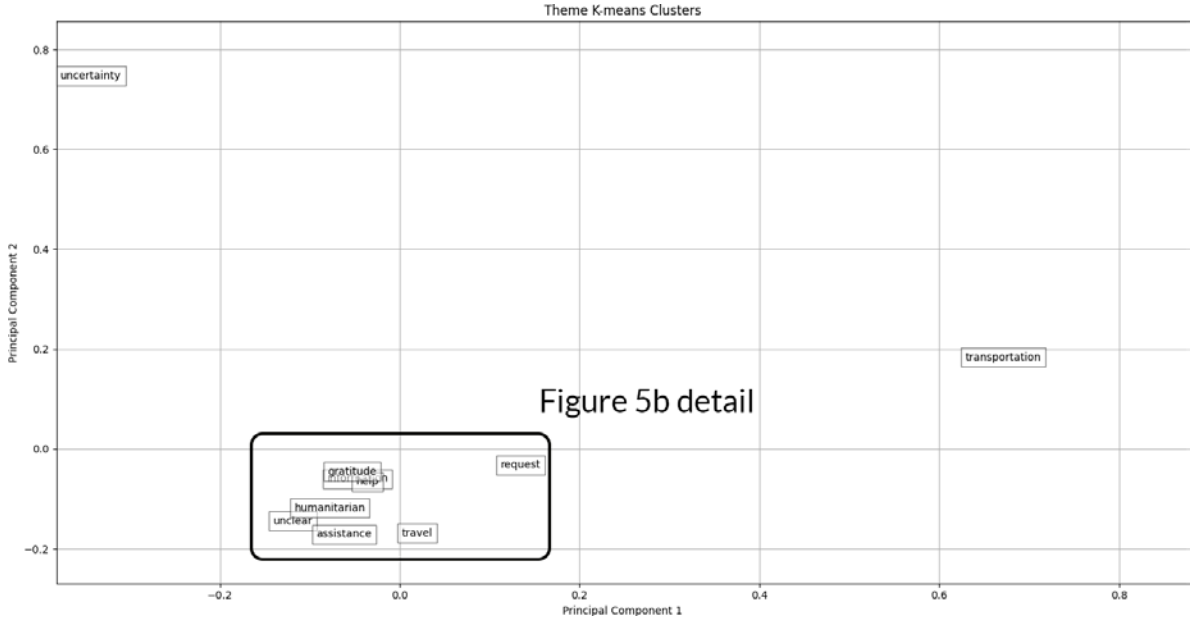


Figure 5a. Theme clusters – overview

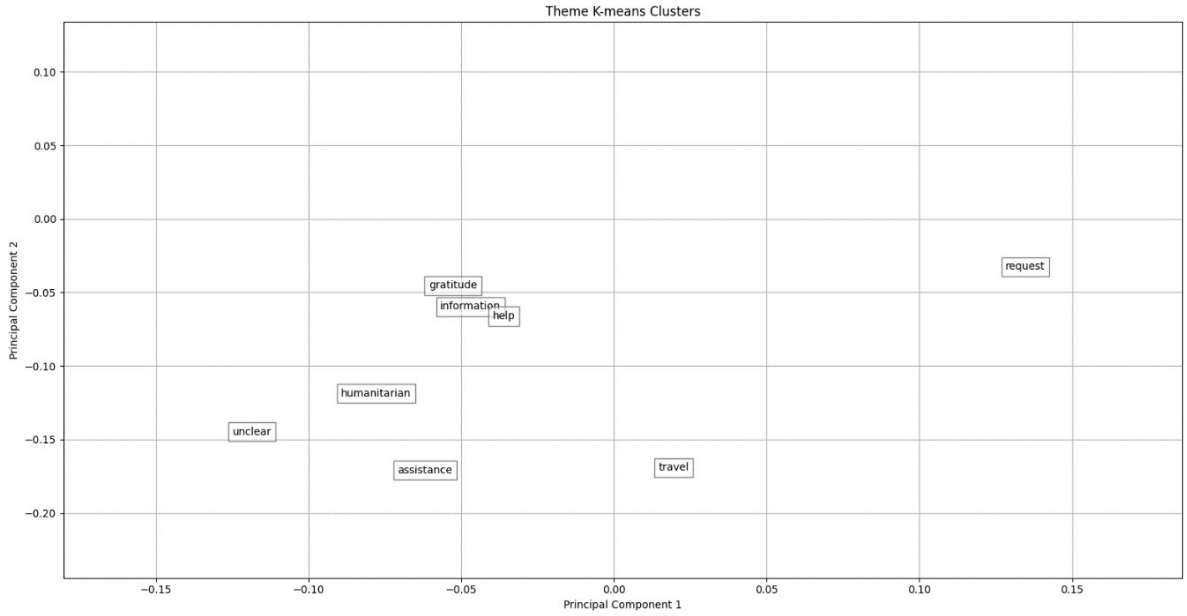


Figure 5b. Theme clusters - detail

Figures 5a and 5b. K-means Theme Clusters



Figure 6. Theme word cloud

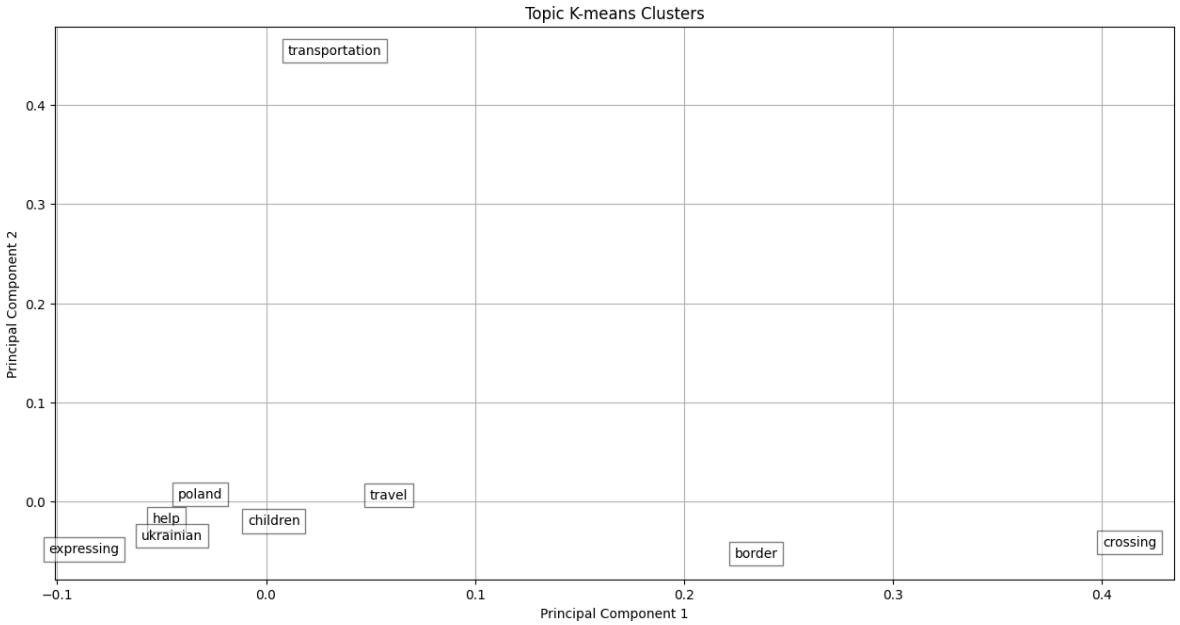


Figure 7. K-means topic clusters

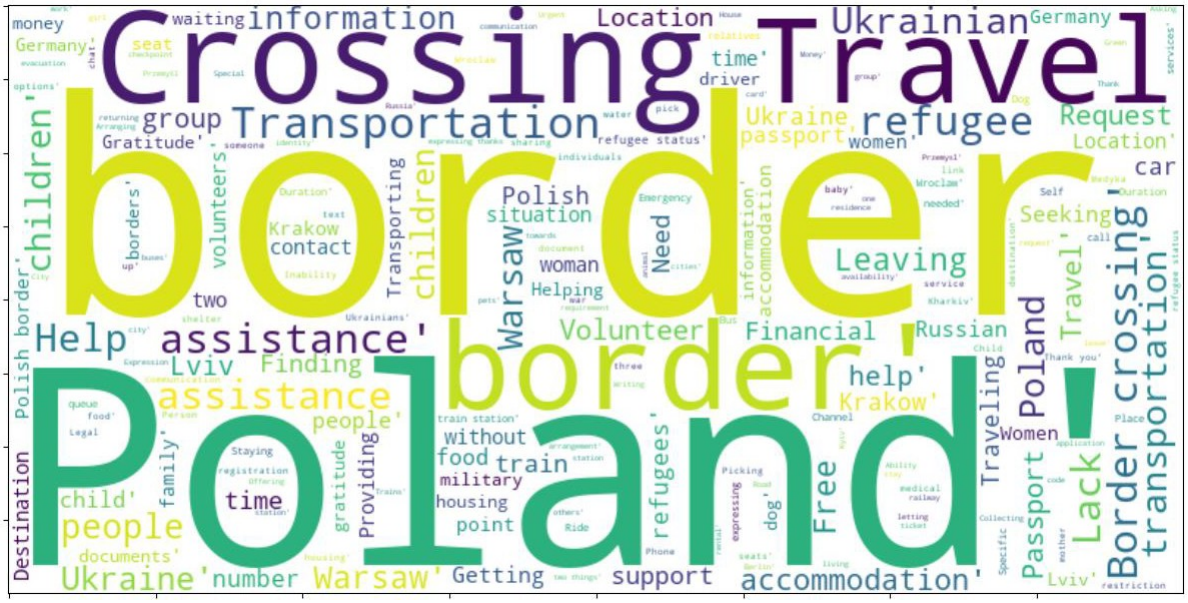


Figure 8. Topic word clouds

Discussion

The LLM generally recognized themes and topics that would be expected when large volumes of people were being forcibly displaced out of Ukraine into Poland. It is interesting to note in the *theme* clustering (Figure 5a), the expected clusters related to supporting people who are looking for transportation and movement in general were found but *uncertainty* also emerged as a cluster. The detail of the theme clusters shown in Figure 5b also shows an interesting pattern of clusters of people who are showing gratitude for information and help. Crossing the border into Poland is clearly the main *topic* of messages that were from the date range analyzed. Again, this is what one would expect. However, it is interesting to note that the topic of *children* emerged as its own cluster as seen in Figure 7. Geography can also be seen within the topic word cloud of Figure 8 such as several specific locations within Poland such as Warsaw, Krakow, Poland/Ukraine border locations such as Lviv and countries such as Germany. These results can reveal potential movement trajectories of the forcibly displaced.

LIMITATIONS AND FUTURE WORK

In the following section we outline several limitations we acknowledge in this work but subsequently provide opportunities for future work. We also identify other general areas for future work.

In this work, we only worked with one version of the chat GPT model (gpt-3.5-turbo-0613). ChatGPT and other LLM are constantly evolving. Therefore, follow-up work could include running the same analysis with a different version of ChatGPT. We also did not compare the NLP functions of ChatGPT against other LLM models such as Google Pa LM and Gemini, Microsoft Copilot, Meta's LLa MA or others. In particular it would be worth investigating how different LLM respond to different types of natural language prompts and the types of results that would be returned.

As per the motivation to investigate the use of LLM specifically to utilize the ability of human natural dialogue to interact with AI, we specifically did not look into using traditional AI natural language processing methods that often require extensive computer science experience to operate that would be beyond the practical scope of operation by most humanitarian professionals.

We only conducted an initial analysis of the text to identify general topics and themes within messages that had been translated to English. This is an area of significant follow-up research. In particular, conducting a sentiment analysis on messages and also identifying erroneous and false information inside of social media messages as per the well-established and recognized phenomena of false information being provided by trolls and malicious actors during a crisis (Federal Emergency Management Agency (FEMA), 2017).

For example, it was pointed out by one of the human evaluators that Russian bots are often active in attempting to spread false information into Telegram social media groups as per statements we found like this *"Stop senseless*

bloodshed! The Russian people are not your enemies, surrender! And you will be treated well: you will receive tasty hot food and warm clothing, and if you are wounded - medical assistance."

We only worked with one message group within one source of social media (Telegram: Help for Ukrainians in Poland). There could be extensive follow-up research on combining multiple streams of social media to develop a more comprehensive picture of situation awareness related to displaced people. Additionally, Telegram has an age bias, most users are younger, more tech-savvy people. Thus, results from this analysis would likely exclude the needs and situations of many older individuals or individuals without smart phones.

We also conducted the topic and thematic analysis based on a translation to English as we acknowledge out limitation and lack of knowledge of Ukrainian and Russian. Deeper incorporation of specialists with expertise in Ukrainian and Russian languages could enable a similar analysis to be done in the source languages and compare those results with those done via English translations. Related to all these points, it is important to acknowledge that ChatGPT itself has gender, racial, and cultural bias derived from its own training data that could impact analysis results (Ray, 2023).

The current word cloud and K-means results presented here are not inherently be useful for humanitarian practitioners. However, future research in automated sorting and filtering of topics, such as messages related to health care and medical needs, would allow for more detailed analysis of specific needs and potential barriers to services. Identifying methods for filtering out "non-relevant" text would also enhance the utility of the word cloud or word frequency analysis. Identifying repeated phrases (i.e. "...schedule a medical appointment..." or "...lon g waiting times...") could also assist in identifying more specific key themes. Thematic analysis using Latent Dirichlet Allocation (LDA) could offer an improved approach over K-means to support this research direction (Westrupp et al., 2022).

Areas of visualization follow-up research would include identification and geocoding locations within the messages to generate geospatial visualizations to understand more specifically about the location of refugees and their movements over space and time, and issues they are facing (Hübl, Cvetojevic, Hochmair, & Paulus, 2017). For example, placing geocoded messages within the context of a map like shown in Figure 1 where official government data sources could be combined with social media data sources to generate a more comprehensive overall picture of the situation of refugees in Poland (Tomaszewski, 2021). There is also extensive visualization research that could be conducted on hierarchal visualization that could be done to allow topic and thematic clusters to be investigated in terms of specific messages within those clusters for more insight and follow-up on specific messages (Elmqvist & Fekete, 2009).

Finally, the work presented here was a pilot study to evaluate the technical aspects of an LLM within the context of a humanitarian crisis situation. Metrics like Kappa and Corpus BLUE scores are easy to understand metrics that can be used to communicate with non-technical humanitarian professionals as an initial step in building trust and acceptance of LLM in the context their ability to be effective at language translations within the context of an international crisis. The ISCRAM community has long made calls for more integration between the academic research and practitioner groups within the overall community. Approaches like one presented here can help begin to cross this divide.

Longer term, it will be important that the approaches like those presented here are evaluated by working humanitarian crisis professionals engaged in situation awareness activities related to forced displacement. For example, analyzingthe accuracy of LLM analysis as compared to existing ground-truthed data sources reflecting the needs and situations of Ukrainian refugees in Poland such as the Multi-Sectoral Needs Assessment (MSNA) report which is based on extensive survey data collection with Ukrainian refugee populations (United Nations High Commissioner for Refugees (UNHCR), 2024).

CONCLUSION

In this work, we evaluated the utility of a large language model (LLM) to develop situation awareness related to the forced displacement of Ukrainian refugees into Poland. Specifically, utilizing text messages derived from the Help for Ukrainians in Poland (UKR: Помощь украинцам в Польше) Telegram message group, we translated messages in a variety of languages including Russian Ukrainian, and Polish into English using ChatGPT. The contents of the messages were then analyzed with ChatGPT to identify topics a nd themes within those messages. The text of topics and themes generated by ChatGPT were then analyzed using K-means clustering to identify the most important topics and themes from within the data that was collected. The language identification and translation capabilities of the LLM were evaluated by two human evaluators and measured using Kappa and Corpus BLUE scorers with the LLM Performing effectively (Kappa: 0.86, Corpus BLUE: 0.46).

It can therefore be concluded that within the scope and acknowledged limitations of the work presented here that the ability of LLM using carefully developed language prompts for large data volume analysis with no need for

manual human analysis shows promise for humanitarian analytics focused on rapidly identifying potential key trends, needs, and locations of displaced people.

As a final takeaway from the work presented in this paper, we asked ChatGPT itself the following for future guidance in this line of work via this prompt:

What would you say to a human academic and practitioner audience interested in information systems for crisis response and management in terms of your ability to help with understanding the situation awareness of Ukrainian refugees who have been forcibly displaced into Poland?

And received back the following reply:

As an AI language model, I can provide valuable assistance in understanding the situation awareness of Ukrainian refugees who have been forcibly displaced into Poland within the context of information systems for crisis response and management.

Here's how I can help:

Data Analysis and Interpretation: I can process and analyze large volumes of data from various sources such as news articles, social media, official reports, and humanitarian organizations' databases to extract insights about the situation of Ukrainian refugees in Poland. This includes understanding their demographics, needs, concerns, and challenges.

Trend Analysis: By analyzing patterns and trends in the data, I can help identify emerging issues, hotspots of activity, shifts in refugee movement, and changes in the humanitarian situation over time. This information is crucial for decision-making and resource allocation by crisis response teams and policymakers.

Risk Assessment: Utilizing natural language processing (NLP) techniques, I can assess the sentiment and tone of refugee communications to identify areas of distress, potential conflicts, or security risks. This aids in proactive risk management and early intervention to prevent escalation of crises.

Resource Allocation Optimization: Through optimization algorithms and predictive modeling, I can assist in determining the most efficient allocation of resources such as food, shelter, medical supplies, and transportation based on real-time needs assessments and demand forecasts.

Communication and Coordination Support: I can facilitate communication and coordination among various stakeholders involved in the crisis response, including government agencies, non-governmental organizations (NGOs), aid workers, volunteers, and affected communities. This ensures timely exchange of critical information and enhances collaboration efforts.

Decision Support Systems: By providing synthesized reports, dashboards, and visualizations, I can support decision-makers in gaining a comprehensive overview of the situation, evaluating the effectiveness of interventions, and making informed decisions to improve crisis response strategies.

Overall, my capabilities extend to assisting academic researchers and practitioners in gaining deeper insights into the situation awareness of Ukrainian refugees in Poland, enabling them to develop more effective information systems and strategies for crisis response and management.

Thus, it is notable that ChatGPT in its current form does have an awareness of how it can support humanitarian operations and it will be interesting to see how these functionalities can be further utilized in the future.

However, LLM are not a panacea. Caution should be considered in that social media and LLM contain bias, falsehoods rumors and other potential pitfalls that can interfere with true understanding of a humanitarian situation (Acerbi & Stubbersfield, 2023; Gallegos et al., 2023). Further evaluation of LLM, social media and combination of official government data sources can show a pathway forward for better understanding of what people who are forcibly displaced are experiencing and how they can be supported.

ACKNOWLEDGMENTS

This research was supported by a grant from the Polish-U.S. Fulbright Commission.

REFERENCES

Acerbi, A., & Stubbersfield, J. M. (2023). Large language models show human-like content biases in transmission chain experiments. *Proceedings of the National Academy of Sciences, 120* (44), e2313790120. doi:10.1073/pnas.2313790120

BBC News. (2022). Ukraine conflict: Simple visual guide to the Russian invasion. Retrieved from https://www.bbc.com/news/world-europe-60506298

Crew, R., Ruetzel, S., Barton, J., Byrnes, T., Wegener, E., Rivera, V., . . . Mushagalusa, J. (2023). How are humanitarians using AI tools like Chat GPT? Retrieved from https://www.calpnetwork.org/blog/how-are-humanitarians-using-ai-tools-like-chat-gpt

Elmqvist, N., & Fekete, J.-D. (2009). Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines. *IEEE transactions on visualization and computer graphics, 16* (3), 439-454. doi:10.1109/TVCG.2009.84

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64. doi:10.1518/001872095779049543

Federal Emergency Management Agency (FEMA). (2017). Harvey Rumor Control. Retrieved from https://www.fema.gov/blog/harvey-rumor-control

Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M. M., Kim, S., Dernoncourt, F., . . . Ahmed, N. K. (2023). Bias and fairness in large language models: A survey. *arXiv preprint arXiv:2309.00770.*

Google Cloud. (n.d.). Understanding the BLEU Score. Retrieved from https://cloud.google.com/translate/automl/docs/evaluate#bleu

gov.pl. (n.d.). Obtain a PESEL Number – a service provided to foreigners. Retrieved from https://www.gov.pl/web/gov/uzyskaj-numer-pesel--usluga-dla-cudzoziemcow-en

Hübl, F., Cvetojevic, S., Hochmair, H., & Paulus, G. (2017). Analyzing refugee migration patterns using geo-tagged tweets. *ISPRS International Journal of Geo-Information, 6*(10), 302. doi:10.3390/ijgi6100302

Hughes, A. L., & Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. *International journal of emergency management, 6* (3-4), 248-260. doi:10.1504/IJEM.2009.031564

International Organization for Migration. (2023a). Poland — FGD Report on Employment, Accommodation, and Future Plans (April 2023). Retrieved from https://dtm.iom.int/reports/poland-fgd-report-employment-accommodation-and-future-plans-april-2023

International Organization for Migration. (2023b). Poland | Displacement Tracking Matrix. Retrieved from https://dtm.iom.int/poland

Kumbhar, R., Mhamane, S., Patil, H., Patil, S., & Kale, S. (2020). *Text document clustering using k-means algorithm with dimension reduction techniques.* Paper presented at the 2020 5th International Conference on Communication and Electronics Systems (ICCES).

OpenAPI. (2024). OpenAI API. Retrieved from https://openai.com/blog/openai-api

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). *Bleu: a method for automatic evaluation of machine translation.* Paper presented at the Proceedings of the 40th annual meeting of the Association for Computational Linguistics.

Power, R., Robinson, B., & Cameron, M. (2022). *Insights from a Decade of Twitter Monitoring for Emergency Management.* Paper presented at the Proceedings of the ISCRAM Asia Pacific Conference.

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*. doi:10.1016/j.iotcps.2023.04.003

Reidy, E. (2022). Is Poland's smooth reception of Ukrainian refugees heading for trouble? Retrieved from https://www.thenewhumanitarian.org/news-feature/2022/08/04/Poland-Ukraine-refugee-concern-grows

Roane, A. (2023). Four ways ChatGPT could help level the humanitarian playing field. Retrieved from https://www.thenewhumanitarian.org/opinion/2023/03/20/ways-chatgpt-could-help-humanitarian-field

Rocca, R., Tamagnone, N., Fekih, S., Contla, X., & Rekabsaz, N. (2023). Natural language processing for humanitarian action: Opportunities, challenges, and the path toward humanitarian NLP. *Frontiers in big Data, 6*, 1082787. doi:10.3389/fdata.2023.1082787

Tanner, L. (2023). *Learning From Innovations Using Artificial Intelligence*. Retrieved from https://humanitariangrandchallenge.org/research-learnings-reports/

Telegram. (2024). Помощь украинцам в Польше (Help for Ukrainians in Poland). Retrieved from t.me/refugeesinPoland

Tomaszewski, B. (2021). *Geographic Information Systems for Disaster Management (Second Edition).* Oxfordshire and New York: Routledge.

UNHCR Regional Bureau for Europe Data Identity Management and Analysis Unit (DIMA). (2022). *Lives on Hold: Profiles and Intentions of Refugees from Ukraine*. Retrieved from

United Nations High Commissioner for Refugees (UNHCR). (2022a). Poland welcomes more than two million refugees from Ukraine. Retrieved from https://www.unhcr.org/news/press/2022/3/6234811a4/poland-welcomes-million-refugees-ukraine.html

United Nations High Commissioner for Refugees (UNHCR). (2022b). Situation Ukraine Refugee Situation. Retrieved from https://data.unhcr.org/en/situations/ukraine

United Nations High Commissioner for Refugees (UNHCR). (2024). *Multi-Sector Needs Assessment - Protection Challenges for Refugees from Ukraine in Poland (MSNA 2023)*. Retrieved from https://data.unhcr.org/en/documents/details/106915

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam Med, 37*(5), 360-363.

Westrupp, E. M., Greenwood, C. J., Fuller-Tyszkiewicz, M., Berkowitz, T. S., Hagg, L., & Youssef, G. (2022). Text mining of Reddit posts: Using latent Dirichlet allocation to identify common parenting issues. *PLoS one, 17*(2), e0262529. doi:10.1371/journal.pone.0262529

Yerramreddy, D. R., Marasani, J., Gowtham, P. S. V., & Abhishek, S. (2023). *An Empirical Analysis of Topic Categorization Using PaLM, GPT and BERT Models.* Paper presented at the 2023 Innovations in Power and Advanced Computing Technologies (i-PACT).

Crisis2Sum: An Exploratory Study on Disaster Summarization from Multiple Streams

Philipp Seeberger
Technische Hochschule Nürnberg
philipp.seeberger@th-nuernberg.de

Korbinian Riedhammer
Technische Hochschule Nürnberg
korbinian.riedhammer@th-nuernberg.de

ABSTRACT

Automatic summarization of natural and human-made disaster events is an important area to increase situational awareness for human response organizations and disaster management. However, the incorporation of multiple data sources poses a challenge to current summarization systems, as the typically large document collections exceed the input limits of neural models. Additionally, Large Language Models (LLM) often omit key information present at different positions in long context inputs. Furthermore, disaster reporting requires fine-grained information content and therefore relaxes the restriction to high compression rates, resulting into rather long summaries. In this work, we study different extractive and LLM-based abstractive baselines and highlight shortcomings in present approaches. Our experimental results on the CrisisFACTS datasets show that LLM-based approaches tend to fail in generating long informative summaries. Taking these limitations into account, we propose a disaster summarization framework and introduce query-focused extensions, which demonstrate advantages and superior performance over the baseline methods.

Keywords

Crisis Informatics, Disaster Summarization, Large Language Models, Integer Linear Programming

INTRODUCTION

Disaster management is a critical process responsible for various key phases (e.g., preparedness, response, and recovery) during mass-emergency incidents, where human decisions heavily rely on a comprehensive understanding of the specific event. Situational awareness aims to provide an ideal understanding of what is happening during an event with numerous actors and sub-events (Vieweg et al., 2010). However, both the lack of information and the overload due to today's information ecosystem can paralyze disaster management and humanitarian response organizations (Imran et al., 2014; McCreadie & Buntain, 2023). It is therefore crucial to automatically localize valuable information and condense it into summary reports that can effectively contribute to situational awareness. Emerging heterogeneous information sources such as social media and microblogging platforms rapidly disseminate crucial details about ongoing events and offer new opportunities to complement traditional approaches (Imran et al., 2014). However, the high-velocity nature of content generation and the properties of different source domains and events pose challenges to current disaster reporting approaches (Kaufhold, 2021; Seeberger & Riedhammer, 2022b).

Unlike most other summarization tasks, we identify several key challenges for disaster reporting: (1) The lack of annotated datasets limits the application of supervised methods and makes it difficult to train sophisticated end-to-end models. (2) Compared with classical query-focused summarization, the information need about an event is complex and requires considering up to dozens of queries simultaneously, depending on the disaster type (McCreadie & Buntain, 2023). For example, it is necessary to capture many aspects such as the burned area, missing persons, contamination level, and evacuation orders during a wildfire event. (3) The multi-stream setting leads to large collections of noisy documents that cause issues regarding the length limitations of recent transformer-based models (Lewis et al., 2020). (4) The target summaries of most summarization tasks are rather short (e.g., up to 250 tokens) and the generation of long summaries (e.g., reports) remains challenging for abstractive methods (Chang et al., 2023; Kryscinski et al., 2022). Recent advances in pre-trained and instruction-following LLMs (Touvron

WiP Paper – Social Media for Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

et al., 2023) as well as distant supervision (Y. Xu & Lapata, 2020) offer opportunities to address these issues; however, they are relatively unexplored in the disaster summarization domain.

In this work, we explore several LLM-based summarization frameworks for generating disaster reports, aiming to highlight new developments across disasters in the form of event timelines (McCreadie & Buntain, 2023). Due to the lack of appropriate training data, we rely on zero-shot and distant supervision methods while focusing on recent open-source LLM developments.

Contributions

(1) Disaster Summarization Framework To address the identified challenges, we design a disaster summarization framework that unifies and extends best practices from different related summarization domains. Automatic evaluation results show improvements over baselines but also highlight shortcomings related to summary generation lengths.

(2) Query-Focused Selection (QILP) We introduce query-focused extensions to an ILP-based extractive summarizer and report absolute improvements of up to 7.65 points for ROUGE-2 w.r.t. NIST reference summaries. In addition, we suggest to use QILP for subsequent selection of abstractive event nuggets and show improvements for the majority of evaluation metrics.

(3) Evaluation & Ablation Study We compare our method with several introduced extractive and abstractive baselines and underscore limitations regarding reference summaries. These are mainly connected to extractive and abstractive biases introduced due to the NIST reference summary creation process. Furthermore, we provide insights into the impact of document sources as well as retrieval pipelines.

RELATED WORK

Query-Focused Summarization

Query-focused summarization aims to generate a summary of given documents that satifies the information need of a specific query. This query often represents keywords, short titles, or long narratives that may cover multiple aspects. Due to the limited availability of query-specific datasets, initial work in this area focused on unsupervised extractive approaches (Litvak & Vanetik, 2017; Wan, 2008; Wan et al., 2007) and later utilize cross-task knowledge from question answering (QA) datasets (Egonmwan et al., 2019; Laskar et al., 2020; Su et al., 2020, 2021; Y. Xu & Lapata, 2020). Another line of research focuses on the utilization of generic summarization tasks by employing pseudo-queries (Vig et al., 2022; Y. Xu & Lapata, 2021) or latent query modeling (Y. Xu & Lapata, 2022). According to Vig et al., most abstractive summarization models follow an extract-then-abstract pipeline and can be categorized into two-step approaches, consisting of an extractor model that scores the source documents and an abstractor model that consumes the documents up to the input length limit (Vig et al., 2022). In contrast, end-to-end approaches aim to address shortcomings due to the performance of the extractor and achieve superior performance for several benchmarks (Vig et al., 2022). However, end-to-end models require appropriate training data and mechanisms that can handle long input texts. Moreover, most approaches focus on tasks with only one or a small number of queries, which is different from the task of summarizing disasters. Our work builds on two-stage models to cope with the lack of annotations as well as noisy and large data streams that are common in disaster events.

Long Document Summarization

The long document summarization field primarily addresses the input length limitations of neural abstractive models. Recent work can be broadly categorized into sparse attention, extract-then-abstract, divide-and-conquer, and hierarchical models (Mao et al., 2022). Sparse attention models such as Longformer (Beltagy et al., 2020), BigBird (Zaheer et al., 2020), and Reformer (Kitaev et al., 2020) utilize different attention mechanisms to reduce the quadratic memory cost of full attention. Several hierarchical models have been proposed for long input processing including models such as HAT-Bart (Rohde et al., 2021) and HMNet (Zhu et al., 2020) which take hierarchical structures into account. Extract-then-abstract methods are similar to the previously mentioned query-focused two-step approach. Most of these two components are trained separately (Bajaj et al., 2021; J. Xu & Durrett, 2019; Y. Zhang et al., 2021), but some work attempts to bridge the two stages by reinforcement learning (Bae et al., 2019; Chen & Bansal, 2018) or joint training (Mao et al., 2022). Divide-and-conquer models break long text inputs into multiple parts, which are summarized separately and then combined into a final summary (Gidiotis & Tsoumakas, 2020; Grail et al., 2021; Y. Zhang et al., 2022). Most recent work perform zero-shot LLM approaches and follow

hierarchical merging and incremental updating strategies (Chang et al., 2023). Hierarchical merging divides the input into several chunks and combines multiple levels of summaries until a desired length (Wu et al., 2021). In contrast, incremental updating aims to overcome the missing context of hierarchical merging by iteratively updating an existing summary with new input chunks (Adams et al., 2023). In our work, we explore divide-and-conquer approaches since they can be easily applied in a zero-shot setting. As previously mentioned, we also rely on the retrieve-then-abstract model. This enables the processing of large and noisy document collections.

Disaster-Related Summarization

Disaster-related methods mostly cover the summarization of social media posts or web news for mass-emergency events. Initial work proposed models for summarizing Twitter posts, which often consist of a classification and subsequent summarization step. The classification step involves the pre-filtering of noise and assignment into predefined categories, which are then summarized into specific or general summaries (P. K. Garg et al., 2023; Rudra et al., 2015, 2018, 2019). Rudra et al. introduced a pipeline consisting of a classifier followed by an extractive summarizer formulated as Integer Linear Programming (ILP) that aims to maximize content words and sub-events coverage w.r.t. specific categories (Rudra et al., 2018). Nguyen and Rudra employs a rationale-motivated classifier to extract rationales that are further used for ILP-based summarization (Nguyen & Rudra, 2022). For temporal summarization, other work focuses on saliency prediction and mechanisms that utilize similarity- or cluster-based methods for summarization (Dusart et al., 2021, 2023; Kedzie et al., 2015). Apart from temporal summarization studies such as TREC TS (Aslam et al., 2015), real-time summarization, with a focus on mobile device notifications and daily email digests from Twitter streams, has gained attention (Sequiera et al., 2018). Most recent work is centered around multi-source datasets, such as CrisisFACTS (McCreadie & Buntain, 2023), and the use of reinforcement learning (Cambrin et al., 2024) or pre-trained LLMs for query-focused disaster reporting (Pereira et al., 2023; Seeberger & Riedhammer, 2022a, 2023). Pereira et al. builds on QA-motivated methods and uses chain-of-thought (CoT) to extract relevant facts from retrieved stream items (Pereira et al., 2023) while FloodBrain (Colvert et al., 2023) aims to generate flood-related reports from web-based content using retrieval-augmented generation. However, classification-based pipelines are limited to predefined categories while summary aspects differ across events. Therefore, we focus on distant supervision for the extractive component. Since summarization on concept level has shown promising results in different domains (Ernst et al., 2022; Riedhammer et al., 2010), we use QA-motivated CoT for fact extraction.

METHOD

In the following, we define our framework for disaster summarization and propose components to overcome the challenges of noisy and large document collections. We give an overview of our approach in Figure 1. In the next section, we formulate our task and present the two-stage pipeline. Then, we detail the Extractor and Abstractor components. Lastly, we introduce extensions to an ILP-based selection model adapted to the query-focused case.

Retrieve-then-Summarize

In disaster summarization, the input consists of $\mathrm{D} = \{d_1, \ldots, d_N\}$ as a set of event-related documents and $\mathrm{Q} = \{q_1, \ldots, q_M\}$ as a set of queries with N and M as the number of documents and queries, respectively. Each query $q \in \mathrm{Q}$ typically consists of a short natural text or a list of keywords and covers a topic or aspect (e.g., wildfire containment). The set of queries Q aims to fulfill the user's information need that is composed of various of these event aspects. Since the initial corpus is often large and still contains irrelevant documents, we employ an Extractor to retrieve and reduce the set of documents to query-related clusters $\mathrm{C} = \{\mathrm{C}^{(q_1)}, \ldots, \mathrm{C}^{(q_M)}\}$ with $\mathrm{C}^{(q_i)} = \{d^{(q_i)}_1, \ldots, d^{(q_i)}_k\}$ consisting of the top-k candidate documents ranked by relevance score. Then, optionally, each cluster $\mathrm{C}^{(q_i)}$ is further refined by a reranking model which operates over all retrieved document-query pairs to estimate semantic matching scores. The Abstractor generates an output summary S (e.g., list of event nuggets) from the cluster pool $\bigcup_{i=1}^{M} \mathrm{C}^{(q_i)}$. Following the view of event timeline summarization, each summary S_t is created w.r.t. a time period $p_t \in \{p_1, \ldots, p_T\}$ with T as the number of time periods and $\mathrm{D}_t \subseteq \mathrm{D}$ as the subset of documents created within the time period p_t. For the sake of clarity, however, we omit the time index t, consider each time period as an independent summarization task and leave redundancy issues for future work.

Extractor

How to model long texts and large document collections is an ongoing research area and various techniques have been proposed to overcome length and GPU memory limitations of neural models by independently processing

chunks. It is common practice to use preprocessing functions such as truncation, chunking, and content selection

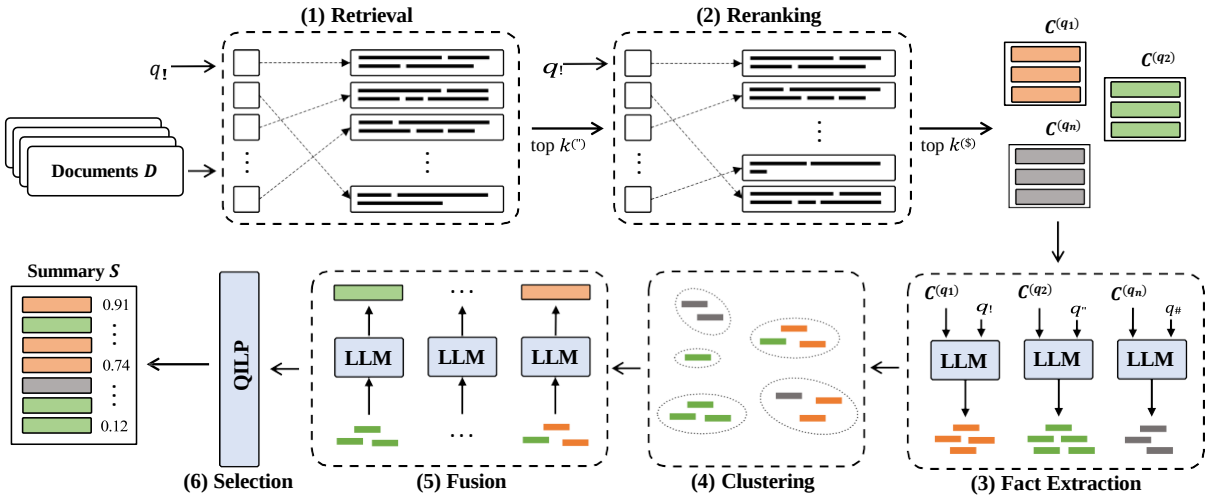


Figure 1. Overview of our framework for disaster summarization. The Extractor consists of stages (1) and (2) while the Abstractor is composed of stages (3), (4), and (5). Stages (1) and (2) retrieve and rerank all documents and extract query-related clusters. (3) generates the query-related facts and (4) clusters the facts into topic groups. Subsequently, (5) fuses the topic clusters to event nuggets. Optionally, (6) selects the event nuggets to create the final summary.

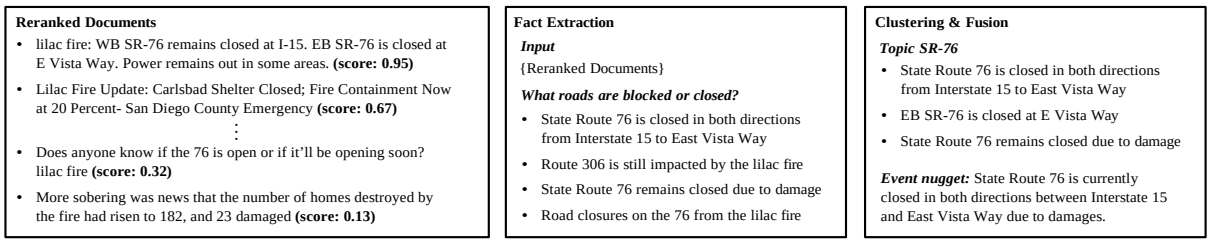


Figure 2. Example for the event nugget creation process w.r.t. the query *What roads are blocked or closed?* and Lilac Wildfire event.

(Dong et al., 2023) while salient content selection can be achieved with retrieval-based models. In our work, we employ content selection methods since text documents from web news, social media, or microblogging platforms are typically short, often irrelevant, and occur in large volumes.

Based on the assumption that salient information occupies only a small portion of long documents (Ding et al., 2020), we employ a two-stage retrieval pipeline as our Extractor in order to reduce the set of documents D to the salient sets of query-related clusters C. As initial step, we retrieve the top-$k^{(1)}$ candidates for each query $q \in Q$ using an efficient sparse retrieval model. Then, we process the concatenation of each query-document pair and obtain a scalar score $s_{ij} = E(q_i, d_j)$, where $E(\cdot)$ represents a neural reranker model and s_{ij} the document relevance
w.r.t. a query q_i. Then, the top-$k^{(2)}$ candidates are selected by descending order of the scores, resulting into a reranked set C'. In Figure 2, we can see that relevant documents achieve higher relevance scores for the query *What roads are blocked or closed?*.

However, the event-related documents may cover different aspects of the event and therefore can obtain multiple high relevance scores for different queries. If we require a single saliency score (e.g., for extractive summarization), we address this issue by a fusion function $f(\cdot)$ to obtain a final saliency score s_j for a document d_j. There exist several rank- and score-based fusion techniques, but in preliminary experiments, we found CombSUM (Fox & Shaw, 1994) to be an effective fusion method. CombSUM calculates the final saliency score for a document as $s_j = \sum_i s_{ij}$ by summing up the relevance scores for each query. As depicted in Figure 1, the Extractor consists

of the main stages (1) Retrieval and (2) Reranking.

Abstractor

In contrast to recent long document processing approaches, we follow a QA-motivated[1] method to enable *event nugget*-based summarization. We first extract all query-relevant facts for the documents and then generate event nuggets employing a clustering-based summarization approach (Ernst et al., 2022). Inspired by (Pereira et al., 2023), we use an instruction-following LLM to generate our facts. Therefore, we utilize the CoT technique and prompt the model to provide a list of facts which serve as evidence for answering the query q based on the provided documents $\{d_1, \ldots, d_j\}$ with $d_i \in C^{(q)}$. The exact prompt is shown in Table 10. For example, we prompt the LLM with the query *What roads are blocked or closed?* and obtain the facts only relevant to this query (Figure 2).

In the next step, all extracted facts are clustered into semantically similar groups. This further prevents redundancy by avoiding multiple facts covering the same piece of information (e.g., *State Route 76 closed*). For clustering, we compute the cosine similarity for each pair of facts using sentence embeddings from a transformer model and then apply affinity propagation. We opt for affinity propagation since the approach does not require specifying the number of clusters and has demonstrated state-of-the-art performance in the disaster summarization domain (Kedzie et al., 2015). Lastly, the resulting documents of the topic clusters are merged with an additional summarization step using the identical LLM (Table 5). Simply selecting the cluster exemplars has shown insufficient results as this approach would discard similar sub-events with different entities such as locations.

For event nugget importance, we recompute the relevance scores for each event nugget and utilize the saliency scores for ranking. That is, each generated event nugget is fed into the neural reranker model, and the resulting scores are used for saliency. The final summary is produced by appending the ranked event nuggets until the desired summary length or by using an optional extractive selection model that we detail in the next section.

Query-Focused Selection

As extractive selection model, we extend previous concept-based methods and model the extractive summarization as an ILP problem (Gillick & Favre, 2009; Riedhammer et al., 2010; Rudra et al., 2018; Seeberger & Riedhammer, 2022a). Concepts typically represent units such as bi-grams, entities, and key-phrases. Our objective is to maximize the coverage of these units while keeping redundancy low. In the rest of this paper, we refer to our extensions as Query-Focused Integer Linear Programming (QILP). The introduced extensions aim to incorporate document saliency and query coverage on the summary level. Note that the extractive summarization can be applied on both the documents directly and generated event nuggets. In particular, we maximize the following objective:

$$\underset{d}{\text{argmax}} \quad \alpha \sum_i w_i c_i + \beta \sum_j s_j d_j - \gamma \sum_p \rho_p (1 - z_p) \tag{1}$$

$$\text{subject to} \quad \forall j \quad \sum_j d_j \leq L \tag{2}$$

$$\forall i \quad \sum_j d_j o_{ij} \geq c_i \tag{3}$$

$$\forall i, j \quad d_j o_{ij} \leq c_i \tag{4}$$

$$\forall_i \, c_i \in \{0,1\} \quad \forall_j \, d_j \in \{0,1\} \quad \forall_k \, z_k \in \{0,1\} \tag{5}$$

Concept Coverage and Document Saliency The first and second term in Equation (1) correspond to the concepts coverage and documents saliency, respectively. Formally, let c_i denote the presence of concept i in the summary, d_j denote the presence of document j in the summary, w_i denote the weight of concept i, and s_j denote the saliency of document j. The final score of the concepts coverage and documents saliency is expressed as the sum of positive weights w_i of concepts and scores s_j of documents covered in the summary. If not otherwise mentioned, we use bi-grams as concepts and frequency as weights.

Consistency Constraints Constraints ensure that if a document is selected, all concepts are also selected and if a concept is selected, at least one document that contains it is selected. Additionally, we ensure that the expected summary length do not exceed the defined length L. Here, the variable o_{ij} denotes the occurrence of concept i in document j. In detail, Equation (2) ensures the length constraint and Equation (3) ensures that if a concept i is selected, then there is at least one summary document covering it. Equation (4) ensures that if a document j is selected to be included in the summary, then all concepts in that document are also included in the objective function.

[1]We follow a typical Question-Answering task and extract the answer (facts) for a question (query).

Query Coverage We assume that only a subset of queries are salient depending on the event and day. Thus, we aim to select a characteristic set of documents which covers the salient queries. We already know that each document j is associated with a relevance score $s_{i\ j}$ w.r.t. to a query q_i. Based upon this, we create a m-dimensional target vector $\tau(\mathrm{D}')$ and summary vector $\pi(S)$ that represents the fraction of documents in D' and S that cover the queries Q. Here, the set D' denotes the generated event nuggets for the abstractive and documents for the extractive case. Specifically, $\tau(\cdot, i)$ and $\pi(\cdot, i)$ denote the fraction of documents that cover query q_i. We compute the target and summary vectors as the normalized mean of relevance scores (Lappas et al., 2012). The third term in Equation
(1) aims to match the target vector τ by imposing penalities on certain assignments with soft constraints. The variable z_p serves as an indicator if the constraint p is satisfied and ρ_p acts as penality score for constraint violation:

$$\forall p \in \mathrm{Q}, \quad z_p = \begin{matrix} 1 & \text{if the query contribution } \pi(S,p) \geq \tau(\mathrm{D}',p) \text{ or} \\ 0 & \text{otherwise.} \end{matrix} \tag{6}$$

In this way, we reward the model for selecting documents which cover salient queries. However, we want to avoid penalization for ignoring queries with low saliency. Therefore, we apply an adaptive truncation threshold θ and only consider the $k^{tr} = \max_{k\ i} \tau(\mathrm{D}', i) \leq \theta$ queries. For instance, the queries q_1 with *What area has the wildfire burned?* and q_2 with *Where are firefighters needed?* are more salient topics w.r.t. a wildfire event than the query
q_3 with *What curfews are in place?*. In this case, we assume the target vector with the corresponding queries q_1, q_2, q_3 is $\tau(\mathrm{D}') = (0.8, 0.6, 0.05)$ and the summary vector should match this target vector. We balance the linear combination of terms in Equation (1) with the hyper-parameters α, β, and γ, which can be any scalar values.

EXPERIMENTAL SETUP

Dataset

For our experiments, we use the CrisisFACTS datasets from 2022 and 2023 (McCreadie & Buntain, 2023). Both datasets include overall 18 disaster events and additional metadata (e.g., search keywords, queries, event types, and source types). Each event-day pair consists of multi-stream data and is extracted from the online sources Twitter, Facebook, Reddit, and Webnews. All documents are provided as stream items. That is, the documents are already segmented into short chunks which we use for our experiments. Along with the muli-stream data, the authors provide ground truth summaries extracted from Wikipedia, ICS-209 forms (Denis et al., 2020), and NIST annotated event nuggets. In Table 4, we provide further details about the number of stream items, days, and event types. More details on the dataset and summary construction process can be found in (McCreadie & Buntain, 2023).

Evaluation

In compliance with (McCreadie & Buntain, 2023), we evaluate the system summaries with ROUGE-1/2/L2 and BERTScore F1-scores (T. Zhang et al., 2020). System summaries for each event-day pair are constructed by the concatenation of all event nuggets (ordered by importance score) up to a threshold of 32 items. For fair comparison, all event nuggets must not exceed 200 characters and each final summary is bounded to $32 * 200 = 6400$ characters. However, we found that 83% (1691 tokens on average) of the NIST reference summaries exceed the token limit 512 of the authors choosen BERTScore model DeBERTa3. In other words, only a fraction of the most important event nuggets is evaluated from a semantic perspective. For completeness, we evaluate and report with BART4 instead but emphasize that the metrics are still flawed by truncation. This model has a context size of 1024 and doubles the event nugget coverage. Other models with larger context size have been reported with no distinguishing power (Bertsch et al., 2023).

Model Details

For the Extractor component, we use BM25 (Robertson & Zaragoza, 2009) with default settings of the *PyTerrier* (Macdonald & Tonellotto, 2020) library and extend it with Bo1 (Amati & Van Rijsbergen, 2002) query expansion to increase the recall. We set the number of feedback terms and documents as 3 and 20, respectively. As reranker, we employ MonoT55 (Nogueira et al., 2020) trained on MS MARCO and select the top-$k^{(1)}$=250 and $k^{(2)}$=50

2We compute all ROUGE scores with the *TorchMetrics* library and use the *PorterStemmer* as preprocessing step.
3https://huggingface.co/microsoft/deberta-xlarge-mnli

4https://huggingface.co/facebook/bart-large-mnli
5https://huggingface.co/castorini/monot5-large-msmarco-10k

candidates for each event-day-query triple. If not otherwise specified, all models and baselines use the same candidate documents as well as relevance and saliency scores.

Regarding the Abstractor component, we choose the *Transformers* (Wolf et al., 2020) library and use MISTRAL-7b[6] (Jiang et al., 2023) with greedy decoding as our instruction-following LLM. To increase the context, we select the left and right documents of each candidate document up to a limit of 500 characters. Due to computational constraints, we limit the number of input tokens for all considered methods to 4096. In terms of clustering, we utilize a *Sentence Transformers* model[7] to compute the sentence embeddings and cosine similarity between fact-pairs. We provide further prompt details in the Appendix Prompt Details.

For QILP, we use bi-grams and frequency as concepts and weights, respectively (Gillick & Favre, 2009; Riedhammer et al., 2010). All scores are normalized and the hyper-parameters are set to $\alpha = 1$, $\beta = 1$, $\gamma = 5$, $\rho = 1$, and $\theta = 0.7$ which we selected based on the TREC TS dataset (Aslam et al., 2015). We refer to our final models as follows: QILP[8] denotes the extractive approach that consists of the Extractor and QILP summarizer. LLMNug-Concat consists of the Extractor and Abstractor based on simple fact concatenation up to the event nugget character limit. That is, we skip the clustering and fusion steps in Figure 1 and use multiple facts to present one event nugget. LLMNug-Concat w/ QILP extends the previous model with QILP for final selection. LLMNug-ClusterFuse employs the proposed fact extraction and cluster fusion while LLMNug-ClusterFuse w/ QILP represents the extension with QILP selection.

Baseline Systems

We aim to establish an upper bound for the comparison of our models as well as extractive and abstractive baselines. Here, we only focus on NIST-based reference summaries since ICS forms are rather long and do not provide any splits for event-day pairs. We generate extractive oracles with a greedy search and denote them as Oracle. Specifically, we start with an empty set and iteratively select documents such that the concatenation maximizes the average of ROUGE-1/2 scores given the reference summary. For feasiblity reasons, we limit the search space to the documents retrieved by our Extractor component. Furthermore, we include a LLaMA-based[9] (Touvron et al., 2023) submission that achieved strong results in the CrisisFACTS 2023 challenge and denote it as LLaMA$_{\text{TREC}}$. LLaMA$_{\text{TREC}}$ similarly prompts query-relevant facts with a QA-motivated approach and concatenates the resulting facts to event nuggets (Seeberger & Riedhammer, 2023). In the following, we introduce the extractive and abstractive baselines.

Extractive We compare with several heuristics and extractive baselines from previous query-focused summarization work. More specifically: Random uses random selection, Lead selects the top document from each query-related cluster, and Greedy selects the overall top ranked documents (McCreadie & Buntain, 2023). As extractive models, we include the query-focused LexRank (Y. Xu & Lapata, 2020), a variant of SubMod (Dasgupta et al., 2013) composed of a linear combination of facility-location, document relevance, dispersion, and cluster contribution and lastly SCC (Rudra et al., 2018) as an ILP-based summarization framework with content word and sub-event coverage maximization. Furthermore, we conduct an ablation study for the different components of our introduced extractive model QILP.

Abstractive We perform experiments with different truncation-based and long document summarization versions while using the identical LLM with zero-shot prompting. For truncated baselines, we include Sum as vanilla summarization, SumCoT with the two-step CoT summarization approach (Wang et al., 2023), and Select+Sum as combination of QILP (L=100) with Sum (Dong et al., 2023). We truncate all inputs w.r.t. the maximal input size of 4096 and ensure that the truncation considers document boundaries. For long document summarization, we include incremental updating and hierarchical merging (Chang et al., 2023). For Incremental updating, we iterate through each chunk of documents while continuously updating a global summary with only new salient information. For Hierarchical merging, we generate summaries for each input chunk and recursively merge the summaries until we have reached one final summary. We restrict these models to the top 500 documents emitted by the Extractor component.

6https://huggingface.co/mistralai/Mistral-7B-Instruct-v0.2
7https://huggingface.co/sentence-transformers/all-mpnet-base-v2
8We use the GLPK 5.0 solver.
9https://huggingface.co/meta-llama/Llama-2-13b-chat-hf

Table 1. Rouge-1/2/L and BERTScore F_1-score (x100) results on NIST reference summaries for CrisisFACTS 2022 (NIST-22) and 2023 (NIST-23). | W | represents the average word count of the system summaries. Bold numbers indicate the best performance whereas underlined numbers denote the best performance for the extractive and abstractive baseline systems.

		ROUGE-1		ROUGE-2		ROUGE-L		BERTScore	
	$\|W\|$	NIST-22	NIST-23	NIST-22	NIST-23	NIST-22	NIST-23	NIST-22	NIST-23
Extractive Baselines									
Random	586	33.00	28.19	9.40	8.29	12.56	11.37	58.99	54.96
Greedy	654	46.42	33.66	24.56	14.11	18.06	13.72	65.06	58.04
Lead	618	42.94	33.23	19.44	12.71	16.59	13.67	64.43	58.40
LexRank	663	45.69	34.35	22.09	13.20	17.40	13.25	64.90	58.42
SubMod	656	43.28	33.61	20.09	12.57	16.65	12.75	64.33	58.21
SCC	867	43.89	36.69	16.14	11.18	15.42	12.79	62.83	57.18
Abstractive Baselines									
Sum	357	33.14	26.30	12.32	8.68	14.04	11.71	58.40	56.06
SumCoT	382	33.80	26.82	12.80	8.39	14.25	11.79	58.26	55.61
Select+Sum	355	32.89	26.19	12.42	8.52	14.18	11.82	58.48	55.82
Incremental	631	39.12	35.24	13.39	11.86	14.72	13.31	59.61	58.95
Hierarchical	555	39.63	30.70	13.12	9.21	15.09	12.31	59.15	56.45
Proposed Approaches									
QILP	710	**47.87**	38.00	**26.07**	**17.49**	**18.21**	14.40	**66.18**	60.52
LLMNug-Concat	717	44.55	38.65	18.59	16.03	16.46	**15.34**	63.87	**62.01**
w/ QILP	759	46.77	**40.14**	19.79	16.48	16.74	15.09	63.82	62.00
LLMNug-ClusterFuse	742	43.99	37.66	15.11	12.05	15.19	13.41	61.89	60.34
w/ QILP	755	44.31	37.95	15.36	12.25	15.34	13.41	62.47	60.48
LLaMA$_{\text{TREC}}$	707	44.95	41.43	18.75	25.10	16.82	25.74	63.02	66.01
Oracle	760	70.37	68.11	63.09	53.54	27.96	29.53	75.46	70.11

MAIN RESULTS

In Table 1, we report the results for NIST reference summaries including the average word count for each generated summary. A general trend is that all proposed methods outperform the extractive and abstractive baseline counterparts. Interestingly, the Greedy heuristic is competitive to our approaches and even surpasses most of the extractive and abstractive baselines. This finding is in line with the results of (McCreadie & Buntain, 2023). The Lead model also shows decent results but the Random selection achieved relatively low performance across all metrics which demonstrates the importance of document saliency (i.e., query-relevance) obtained by the Extractor.

Proposed Models The extractive summaries of our proposed approach QILP achieve better results in comparison to all extractive and abstractive baselines, especially for ROUGE-2 (26.07 and 17.49). Surprisingly, the simple concatenation of generated facts outperforms the computational more expensive clustering-based summarization approach. This highlights that disaster reports are rather focused on informativeness than coherence and fluency. The application of extractive selection also improves the LLM-based models for the majority of reported metrics. Critically, we find that most of the best performing models are extractive for NIST-22 and abstractive for NIST-23. This suggests that the reference summaries are potentially biased towards extractive and abstractive systems which could be caused by the pooling-based annotation process. For example, LLaMA$_{\text{TREC}}$ was part of the pooling-based evaluation of CrisisFACTS 2023 and therefore shows the best results for NIST-23. This finding also supports the assumption of biased reference summaries. Regarding the extractive Oracle, we still can see large gaps in terms of ROUGE-1/2/L and BERTScore. These results depict the necessity for further developments in the disaster summarization domain.

ICS-209 and Wikipedia We present the evaluation results for ICS-209 and Wikipedia summaries in Table 2. Overall, we find that the system ranks differ from the NIST reference summaries. Shorter and more general summaries achieve the best results for Wikipedia while longer and more detailed summaries have the best coverage for ICS-209 forms. This is reasonable for this kind of summary formats, where Wikipedia articles provide high-level summaries for a public audience and ICS-209 forms represent informative reports written by emergency-response personnel (McCreadie & Buntain, 2023). This suggests the selection of summarization systems tailored to the

Table 2. Rouge-1/2/L and BERTScore F_1-score (x100) results on ICS-209 forms (ICS-22) and Wikipedia (WIKI) reference summaries. ICS summaries are only available for CrisisFACTS 2022. Bold numbers indicate the best performance. Due to context size limitations, we discard the BERTScore for ICS-22.

	ROUGE-1		ROUGE-2		ROUGE-L		BERTScore	
	ICS-22	WIKI	ICS-22	WIKI	ICS-22	WIKI	ICS-22	WIKI
Greedy	27.98	8.85	6.24	2.52	9.48	4.58	-	48.86
SCC	28.34	7.79	5.42	2.51	9.21	3.91	-	49.32
Incremental	28.34	9.85	6.33	3.37	9.87	5.19	-	53.32
Hierarchical	28.06	**10.34**	6.31	**3.61**	10.41	**5.57**	-	**53.71**
QILP	28.22	8.61	6.24	2.90	9.43	4.44	-	49.87
LLMNug-Concat	29.60	8.26	6.92	2.85	10.23	4.48	-	51.87
w/ QILP	30.38	7.96	**7.09**	2.71	10.30	4.29	-	51.35
LLMNug-ClusterFuse	30.69	7.99	7.08	2.57	**10.61**	4.37	-	51.50
w/ QILP	**30.74**	8.06	7.06	2.74	10.57	4.40	-	51.87

Table 3. Rouge-1/2/L and BERTScore F_1-score (x100) results for the ablation study. We removed or substituted the components of our model and report the absolute increase (↑) or decrease (↓) w.r.t. NIST reference summaries.

	ROUGE-1		ROUGE-2		ROUGE-L		BERTScore	
	NIST-22	NIST-23	NIST-22	NIST-23	NIST-22	NIST-23	NIST-22	NIST-23
QILP	47.87	38.00	26.07	17.49	18.21	14.40	66.18	60.52
w/o Q-Cov	↓0.14	↓0.14	↓0.17	↓0.09	↓0.03	↑0.03	↓0.15	↓0.14
w/o Sal	↓5,74	↓0,75	↓7,54	↓1,76	↓2,57	↓0,25	↓2,06	↓0,83
w/o Sal+Q-Cov	↓5.98	↓1.42	↓7.65	↓2.21	↓2.37	↓0.59	↓2.12	↓1.08
w/ Entities	↓0,06	↓2,67	↓0,65	↓2,23	↑0,17	↓0,52	↓0,19	↓1,26
w/o Reranking	↓11,81	↓16,74	↓10,13	↓11,11	↓3,10	↓4,81	↓6,80	↓8,42

end-user and information need. However, we emphasize that all evaluated summarization models are used in the unsupervised and zero-shot setting. Appropriate datasets and model fine-tuning can further improve the results but data collections are costly to obtain in the disaster summarization domain.

Are we gaming the benchmark? If we consider summary lengths, one can easily spot the correlating ROUGE-1/2/L and BERTScore scores. Here, the question arises if simply producing long summaries can be seen as *gaming the benchmark*. In fact, we measure a moderate Spearman's rank correlation of 0.43 (0.48) for ROUGE-2 (BERTScore) and note that length bias in neural summarization (Guo & Vosoughi, 2023; Sun et al., 2019) as well as long document summarization (Chang et al., 2023) evaluation is an active research area. However, we argue that producing long and detailed summaries is a desirable feature in the field of disaster reporting. Therefore, we see our event nugget-based summarization as an extension to obtain more informative summaries, but would also like to emphasize the length limitations of our abstractive baselines. In Appendix Examples, we show excerpts for summaries about the *Kincade Wildfire 2019* event.

FURTHER ANALYSIS

In addition to the main experiments, we further assess the impact of different setups and components. Therefore, we analyse the influence of retrieval pipelines, document sources, and provide an ablation study for the QILP model components.

Influence of Retrieval Figure 3 presents the ROUGE-2 scores for different retrieval pipelines. We conducted experiments for the sparse retrieval models DFReeKLIM (Amati et al., 2011) and BM25, sparse neural search model SPLADE10 (Formal et al., 2022), and rerankers ASNQ11 (S. Garg et al., 2020) and MonoT5 with BM25 as the first-stage retriever. The retrieval experiments highlight two important points. First, the usage of reranking

10https://huggingface.co/naver/splade-cocondenser-ensembledistil
11TANDA RoBERTa-Large ASNQ-Wiki-QA

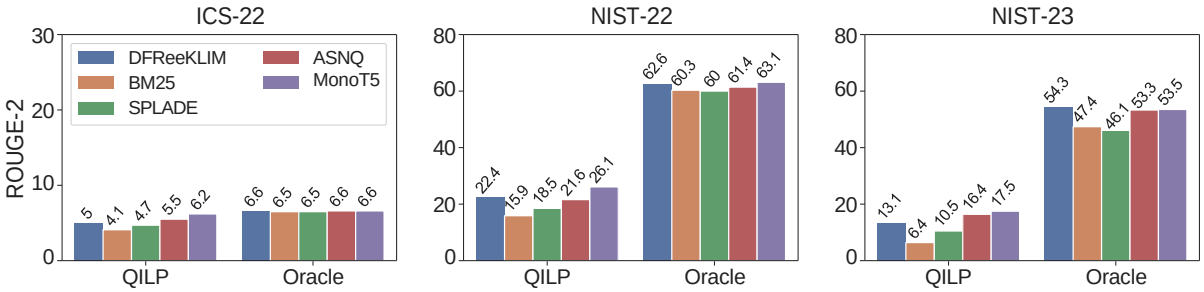


Figure 3. Rouge-2 F_1-score (x100) results for the different retrieval pipelines. We report the performance for QILP and the extractive Oracle. The rerankers ASNQ and MonoT5 use BM25 for first-stage retrieval.

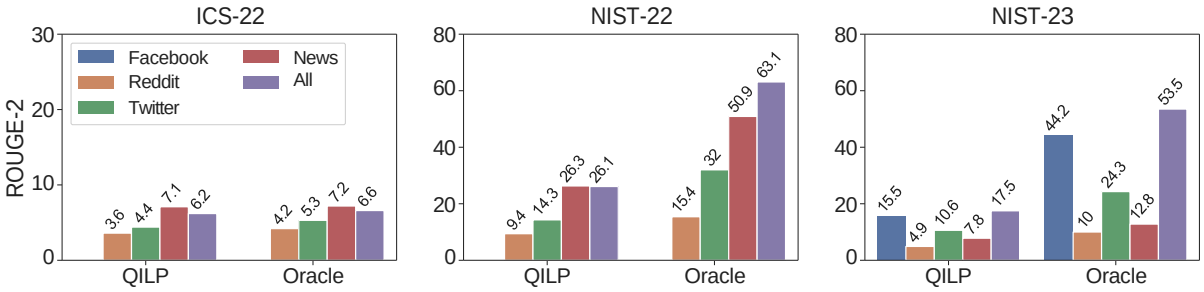


Figure 4. Rouge-2 F_1-score (x100) results for the different available document sources. We report the performance for QILP and the extractive Oracle. The label *All* denotes the mixture of all document sources.

significantly improves the results over the first-stage retrieval and highlights the importance of appropriate document scoring w.r.t. the downstream task. Second, dedicated efficient sparse retrieval models like DFReeKLIM can outperform off-the-shelf models such as BM25 by large margins but still lack behind the two-stage approaches. However, the DFReeKLIM results can be strongly biased due to the pooling-based process as the organizers use two basic DFReeKLIM variants as baseline systems.

Influence of Source Figure 4 presents the ROUGE-2 scores by restricting the input documents to the source types. Note that we only had access to Facebook posts for CrisisFACTS 2023 which explains the missing scores for ICS-22 and NIST-22 summaries. First, we can see that posts from Reddit and Twitter are mostly outperformed by the sources Facebook, News, and All for both QILP and Oracle. We assume that News already summarizes important facts about an event and Facebook often refers to published news articles or include posts written by news accounts (McCreadie & Buntain, 2023). In contrast, Twitter posts contain much noise and Reddit posts are only available in smaller scale. The drop of News for NIST-23 can be explained by both the low number of available stream items and the possible bias towards Facebook posts. This highlights the benefit of multiple document sources as long as the availability of traditional sources is limited.

Influence of Components Table 3 presents the results of our component ablation for the best performing QILP model. As already mentioned in previous sections, removing the reranking component yields the largest drop for all metrics. Exchanging the Bi-gram concepts with Entity concepts leads to an further decrease in performance that mirrors the effectiveness of simple Bi-grams for content coverage. For our saliency (Sal) and query coverage (Q-Cov) extensions, we found that removing saliency leads to a higher performance drop while the largest decrease occurs if we remove both components. In other words, the saliency and query coverage terms can be seen as complementary and not just as substitute. However, we only consider retrieval relevance scores as document saliency and would expect a performance increase by incorporating more sophisticated saliency predictors (Kedzie et al., 2015).

Efficiency Practical utilization of information systems in emergency and disaster response necessitates the deployment of efficient real-time systems. Therefore, we analyze the computation time of the proposed components concerning the event *Lilac Wildfire 2017*. All experiments are conducted with AMD EPYC 7662 CPUs and one Nvidia A100 GPU. On average, the stages of Retrieval and Reranking require 1.2 and 67.4 seconds for each

event-day pair and provide the input for the subsequent summarization components. While the extractive approach QILP only takes 0.8 seconds, the abstractive LLMNug-Concat and LLMNug-ClusterFuse approaches require 673.2 and 839.7 seconds, respectively. This highlights the computational overhead of the incorporated LLMs, which may be insufficient for summarization updates in shorter intervals. Using proprietary LLMs via API access can enhance performance and avoid the need for dedicated infrastructure. However, this also introduces latency issues and may raise data privacy concerns.

CONCLUSION

In this work, we studied the task of disaster summarization from multiple streams. We establish several unsupervised and zero-shot baselines and assess the performance with CrisisFACTS, a state-of-the-art dataset for disaster summarization. Current summarization systems reveal shortcomings related to large document collections and face challenges in generating long disaster reports. To address these issues, we introduce a retrieve-then-summarize approach and extend previous research from the query-focused and long document summarization domain. Experimental results show that our approach outperforms baselines, but we also conclude that simple extractive methods achieve competitive or better results. In addition, we examine the influence of different components and point out shortcomings related to summary length that make fair evaluation difficult. Depending on the dataset, we find that current reference summaries favor either extractive or abstractive techniques. This suggests that further efforts are needed to develop datasets and automatic evaluation metrics which take bias issues into account. Interesting future directions include the automatic evaluation of disaster reports, comprehensive human evaluation, and benchmarking of larger (proprietary) LLMs.

REFERENCES

Adams, G., Fabbri, A., Ladhak, F., Lehman, E., & Elhadad, N. (2023, December). From Sparse to Dense: GPT-4 Summarization with Chain of Density Prompting. In Y. Dong, W. Xiao, L. Wang, F. Liu, & G. Carenini (Eds.), *Proceedings of the 4th New Frontiers in Summarization Workshop* (pp. 68–74). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.newsum-1.7

Amati, G., Amodeo, G., Bianchi, M., Celi, A., Nicola, C., Flammini, M., Gaibisso, C., Gambosi, G., & Marcone, G. (2011). Fub, iasi-cnr, univaq at trec 2011 microblog track. *The Twentieth Text REtrieval Conference (TREC 2011) Proceedings*.

Amati, G., & Van Rijsbergen, C. J. (2002). Probabilistic models of information retrieval based on measuring the divergence from randomness. *ACM Transactions on Information Systems*, *20*(4), 357–389. https://doi.org/10.1145/582415.582416

Aslam, J., Diaz, F., Ekstrand-Abueg, M., McCreadie, R., Pavlu, V., & Sakai, T. (2015). Trec 2015 temporal summarization track overview. *The Twenty-Fourth Text REtrieval Conference, TREC 2015, November*.

Bae, S., Kim, T., Kim, J., & Lee, S.-g. (2019, November). Summary Level Training of Sentence Rewriting for Abstractive Summarization. In L. Wang, J. C. K. Cheung, G. Carenini, & F. Liu (Eds.), *Proceedings of the 2nd Workshop on New Frontiers in Summarization* (pp. 10–20). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-5402

Bajaj, A., Dangati, P., Krishna, K., Ashok Kumar, P., Uppaal, R., Windsor, B., Brenner, E., Dotterrer, D., Das, R., & McCallum, A. (2021, August). Long Document Summarization in a Low Resource Setting using Pretrained Language Models. In J. Kabbara, H. Lin, A. Paullada, & J. Vamvas (Eds.), *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Student Research Workshop* (pp. 71–80). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-srw.7

Beltagy, I., Peters, M. E., & Cohan, A. (2020). Longformer: The Long-Document Transformer. *arXiv:2004.05150*.

Bertsch, A., Alon, U., Neubig, G., & Gormley, M. R. (2023). Unlimiformer: Long-range transformers with unlimited length input. *37th Conference on Neural Information Processing Systems*.

Cambrin, D. R., Cagliero, L., & Garza, P. (2024). Dqnc2s: Dqn-based cross-stream crisis event summarizer.

Chang, Y., Lo, K., Goyal, T., & Iyyer, M. (2023, October). BooookScore: A systematic exploration of book-length summarization in the era of LLMs.

Chen, Y.-C., & Bansal, M. (2018, July). Fast Abstractive Summarization with Reinforce-Selected Sentence Rewriting. In I. Gurevych & Y. Miyao (Eds.), *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 675–686). Association for Computational Linguistics. https://doi.org/10.18653/v1/P18-1063

Colvert, G., Silverberg, L., Darm, P., & Kasmanoff, N. (2023). Floodbrain: Flood disaster reporting by web-based retrieval augmented generation with an llm. *6th Workshop on Artificial Intelligence for Humanitarian Assistance and Disaster Response*.

Dasgupta, A., Kumar, R., & Ravi, S. (2013, August). Summarization Through Submodularity and Dispersion. In H. Schuetze, P. Fung, & M. Poesio (Eds.), *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1014–1022). Association for Computational Linguistics.

Denis, L. S., Mietkiewicz, N., Short, K., Buckland, M., & Balch, J. (2020). ICS-209-PLUS - An all-hazards dataset mined from the US National Incident Management System 1999-2014. https://doi.org/10.6084/m9.figshare.8048252.v14

Ding, M., Zhou, C., Yang, H., & Tang, J. (2020). Cogltx: Applying bert to long texts. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, & H. Lin (Eds.), *Advances in neural information processing systems* (pp. 12792–12804, Vol. 33). Curran Associates, Inc.

Dong, Z., Tang, T., Li, L., & Zhao, W. X. (2023, February). A Survey on Long Text Modeling with Transformers.

Dusart, A., Pinel-Sauvagnat, K., & Hubert, G. (2021). ISSumSet: A tweet summarization dataset hidden in a TREC track. *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, 665–671. https://doi.org/10.1145/3412841.3441946

Dusart, A., Pinel-Sauvagnat, K., & Hubert, G. (2023). TSSuBERT: How to Sum Up Multiple Years of Reading in a Few Tweets. *ACM Transactions on Information Systems*, *41*(4), 1–33. https://doi.org/10.1145/3581786

Egonmwan, E., Castelli, V., & Sultan, M. A. (2019, November). Cross-Task Knowledge Transfer for Query-Based Text Summarization. In A. Fisch, A. Talmor, R. Jia, M. Seo, E. Choi, & D. Chen (Eds.), *Proceedings of the 2nd Workshop on Machine Reading for Question Answering* (pp. 72–77). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-5810

Ernst, O., Caciularu, A., Shapira, O., Pasunuru, R., Bansal, M., Goldberger, J., & Dagan, I. (2022, July). Proposition-level clustering for multi-document summarization. In M. Carpuat, M.-C. de Marneffe, & I. V. Meza Ruiz (Eds.), *Proceedings of the 2022 conference of the north american chapter of the association for computational linguistics: Human language technologies* (pp. 1765–1779). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.naacl-main.128

Formal, T., Lassance, C., Piwowarski, B., & Clinchant, S. (2022). From distillation to hard negative sampling: Making sparse neural ir models more effective. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2353–2359. https://doi.org/10.1145/3477495.3531857

Fox, E. A., & Shaw, J. A. (1994). Combination of multiple searches. *Proceedings of TREC-2*.

Garg, P. K., Chakraborty, R., & Dandapat, S. K. (2023). OntoDSumm: Ontology-Based Tweet Summarization for Disaster Events. *IEEE Transactions on Computational Social Systems*, 1–16. https://doi.org/10.1109/TCSS.2023.3266025

Garg, S., Vu, T., & Moschitti, A. (2020). Tanda: Transfer and adapt pre-trained transformer models for answer sentence selection. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(05), 7780–7788. https://doi.org/10.1609/aaai.v34i05.6282

Gidiotis, A., & Tsoumakas, G. (2020). A Divide-and-Conquer Approach to the Summarization of Long Documents. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *28*, 3029–3040. https://doi.org/10.1109/TASLP.2020.3037401

Gillick, D., & Favre, B. (2009). A Scalable Global Model for Summarization. *Proceedings of the Workshop on Integer Linear Programming for Natural Language Processing*, 10–18.

Grail, Q., Perez, J., & Gaussier, E. (2021, April). Globalizing BERT-based Transformer Architectures for Long Document Summarization. In P. Merlo, J. Tiedemann, & R. Tsarfaty (Eds.), *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume* (pp. 1792–1810). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.eacl-main.154

Guo, X., & Vosoughi, S. (2023, December). Length does matter: Summary length can bias summarization metrics. In H. Bouamor, J. Pino, & K. Bali (Eds.), *Proceedings of the 2023 conference on empirical methods in natural language processing* (pp. 15869–15879). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.984

Imran, M., Castillo, C., Lucas, J., Meier, P., & Vieweg, S. (2014). Aidr: Artificial intelligence for disaster response. *Proceedings of the 23rd International Conference on World Wide Web*, 159–162. https://doi.org/10.1145/2567948.2577034

Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D. S., de las Casas, D., Bressand, F., Lengyel, G., Lample, G., Saulnier, L., Lavaud, L. R., Lachaux, M.-A., Stock, P., Scao, T. L., Lavril, T., Wang, T., Lacroix, T., & Sayed, W. E. (2023). Mistral 7b.

Kaufhold, M.-A. (2021). *Information Refinement Technologies for Crisis Informatics: User Expectations and Design Principles for Social Media and Mobile Apps*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-33341-6

Kedzie, C., McKeown, K., & Diaz, F. (2015, July). Predicting Salient Updates for Disaster Summarization. In C. Zong & M. Strube (Eds.), *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 1608–1617). Association for Computational Linguistics. https://doi.org/10.3115/v1/P15-1155

Kitaev, N., Kaiser, L., & Levskaya, A. (2020). Reformer: The Efficient Transformer. *International Conference on Learning Representations*.

Kryscinski, W., Rajani, N., Agarwal, D., Xiong, C., & Radev, D. (2022, December). BOOKSUM: A collection of datasets for long-form narrative summarization. In Y. Goldberg, Z. Kozareva, & Y. Zhang (Eds.), *Findings of the association for computational linguistics: Emnlp 2022* (pp. 6536–6558). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.findings-emnlp.488

Lappas, T., Crovella, M., & Terzi, E. (2012). Selecting a characteristic set of reviews. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 832–840. https://doi.org/10.1145/2339530.2339663

Laskar, M. T. R., Hoque, E., & Huang, J. X. (2020, December). WSL-DS: Weakly Supervised Learning with Distant Supervision for Query Focused Multi-Document Abstractive Summarization. In D. Scott, N. Bel, & C. Zong (Eds.), *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 5647–5654). International Committee on Computational Linguistics. https://doi.org/10.18653/v1/2020.coling-main.495

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2020, July). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 7871–7880). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.703

Litvak, M., & Vanetik, N. (2017, April). Query-based summarization using MDL principle. In G. Giannakopoulos, E. Lloret, J. M. Conroy, J. Steinberger, M. Litvak, P. Rankel, & B. Favre (Eds.), *Proceedings of the MultiLing 2017 Workshop on Summarization and Summary Evaluation Across Source Types and Genres* (pp. 22–31). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-1004

Macdonald, C., & Tonellotto, N. (2020). Declarative Experimentation in Information Retrieval Using PyTerrier. *Proceedings of the 2020 ACM SIGIR on International Conference on Theory of Information Retrieval*, 161–168. https://doi.org/10.1145/3409256.3409829

Mao, Z., Wu, C. H., Ni, A., Zhang, Y., Zhang, R., Yu, T., Deb, B., Zhu, C., Awadallah, A., & Radev, D. (2022, May). DYLE: Dynamic Latent Extraction for Abstractive Long-Input Summarization. In S. Muresan, P. Nakov, & A. Villavicencio (Eds.), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1687–1698). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.118

McCreadie, R., & Buntain, C. L. (2023). CrisisFACTS: Buidling and Evaluating Crisis Timelines. *Proceedings of the 20th International ISCRAM Conference*, 320–339. https://doi.org/10.59297/JVQZ9405

Nguyen, T. H., & Rudra, K. (2022). Towards an Interpretable Approach to Classify and Summarize Crisis Events from Microblogs. *Proceedings of the ACM Web Conference 2022*, 3641–3650. https://doi.org/10.1145/3485447.3512259

Nogueira, R., Jiang, Z., Pradeep, R., & Lin, J. (2020, November). Document ranking with a pretrained sequence-to-sequence model. In T. Cohn, Y. He, & Y. Liu (Eds.), *Findings of the association for computational linguistics: Emnlp 2020* (pp. 708–718). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.findings-emnlp.63

Pereira, J., Fidalgo, R., Lotufo, R., & Nogueira, R. (2023). Crisis Event Social Media Summarization with GPT-3 and Neural Reranking. *Proceedings of the 20th International ISCRAM Conference*, 371–384. https://doi.org/10.59297/JJYT4136

Riedhammer, K., Favre, B., & Hakkani-Tu¨r, D. (2010). Long story short – Global unsupervised models for keyphrase based meeting summarization. *Speech Communication*, *52*(10), 801–815. https://doi.org/10.1016/j.specom.2010.06.002

Robertson, S., & Zaragoza, H. (2009). The probabilistic relevance framework: Bm25 and beyond. *Found. Trends Inf. Retr.*, *3*(4), 333–389. https://doi.org/10.1561/1500000019

Rohde, T., Wu, X., & Liu, Y. (2021, September). Hierarchical Learning for Generation with Long Source Sequences.

Rudra, K., Ghosh, S., Ganguly, N., Goyal, P., & Ghosh, S. (2015). Extracting Situational Information from Microblogs during Disaster Events: A Classification-Summarization Approach. *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, 583–592. https://doi.org/10.1145/2806416.2806485

Rudra, K., Goyal, P., Ganguly, N., Imran, M., & Mitra, P. (2019). Summarizing Situational Tweets in Crisis Scenarios: An Extractive-Abstractive Approach. *IEEE Transactions on Computational Social Systems*, *6*(5), 981–993. https://doi.org/10.1109/TCSS.2019.2937899

Rudra, K., Goyal, P., Ganguly, N., Mitra, P., & Imran, M. (2018). Identifying Sub-events and Summarizing Disaster-Related Information from Microblogs. *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 265–274. https://doi.org/10.1145/3209978.3210030

Seeberger, P., & Riedhammer, K. (2022a). Combining deep neural reranking and unsupervised extraction for multi-query focused summarization. In I. Soboroff & A. Ellis (Eds.), *Proceedings of the thirty-first text retrieval conference, TREC 2022, online, november 15-19, 2022* (Vol. 500-338). National Institute of Standards; Technology (NIST).

Seeberger, P., & Riedhammer, K. (2022b). Enhancing crisis-related tweet classification with entity-masked language modeling and multi-task learning. *Proceedings of the Second Workshop on NLP for Positive Impact (NLP4PI)*, 70–78.

Seeberger, P., & Riedhammer, K. (2023). Multi-query focused disaster summarization via instruction-based prompting. In I. Soboroff (Ed.), *Proceedings of the thirty-second text retrieval conference, TREC 2023, gaithersburg, november 13-17, 2023*. National Institute of Standards; Technology (NIST).

Sequiera, R., Tan, L., & Lin, J. (2018). Overview of the trec 2018 real-time summarization track. *The Twenty-Seventh Text REtrieval Conference, TREC 2018, November*.

Su, D., Xu, Y., Yu, T., Siddique, F. B., Barezi, E., & Fung, P. (2020, December). CAiRE-COVID: A Question Answering and Query-focused Multi-Document Summarization System for COVID-19 Scholarly Information Management. In K. Verspoor, K. B. Cohen, M. Conway, B. de Bruijn, M. Dredze, R. Mihalcea, & B. Wallace (Eds.), *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.nlpcovid19-2.14

Su, D., Yu, T., & Fung, P. (2021, August). Improve Query Focused Abstractive Summarization by Incorporating Answer Relevance. In C. Zong, F. Xia, W. Li, & R. Navigli (Eds.), *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021* (pp. 3124–3131). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.findings-acl.275

Sun, S., Shapira, O., Dagan, I., & Nenkova, A. (2019, June). How to compare summarizers without target length? pitfalls, solutions and re-examination of the neural summarization literature. In A. Bosselut, A. Celikyilmaz, M. Ghazvininejad, S. Iyer, U. Khandelwal, H. Rashkin, & T. Wolf (Eds.), *Proceedings of the workshop on methods for optimizing and evaluating neural language generation* (pp. 21–29). Association for Computational Linguistics. https://doi.org/10.18653/v1/W19-2303

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozie`re, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). Llama: Open and efficient foundation language models.

Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events: What twitter may contribute to situational awareness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1079–1088. https://doi.org/10.1145/1753326.1753486

Vig, J., Fabbri, A., Kryscinski, W., Wu, C.-S., & Liu, W. (2022, July). Exploring Neural Models for Query-Focused Summarization. In M. Carpuat, M.-C. de Marneffe, & I. V. Meza Ruiz (Eds.), *Findings of the Association for Computational Linguistics: NAACL 2022* (pp. 1455–1468). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.findings-naacl.109

Wan, X. (2008). Using only cross-document relationships for both generic and topic-focused multi-document summarizations. *Information Retrieval*, *11*(1), 25–49. https://doi.org/10.1007/s10791-007-9037-5

Wan, X., Yang, J., & Xiao, J. (2007). Manifold-ranking based topic-focused multi-document summarization. *Proceedings of the 20th International Joint Conference on Artifical Intelligence*, 2903–2908.

Wang, Y., Zhang, Z., & Wang, R. (2023, July). Element-aware summarization with large language models: Expert-aligned evaluation and chain-of-thought method. In A. Rogers, J. Boyd-Graber, & N. Okazaki (Eds.), *Proceedings of the 61st annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 8640–8665). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.acl-long.482

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., . . . Rush, A. (2020). Transformers: State-of-the-art natural language processing. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 38–45. https://doi.org/10.18653/v1/2020.emnlp-demos.6

Wu, J., Ouyang, L., Ziegler, D. M., Stiennon, N., Lowe, R., Leike, J., & Christiano, P. (2021, September). Recursively Summarizing Books with Human Feedback.

Xu, J., & Durrett, G. (2019, November). Neural Extractive Text Summarization with Syntactic Compression. In K. Inui, J. Jiang, V. Ng, & X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3292–3303). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1324

Xu, Y., & Lapata, M. (2020, November). Coarse-to-Fine Query Focused Multi-Document Summarization. In B. Webber, T. Cohn, Y. He, & Y. Liu (Eds.), *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 3632–3645). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.296

Xu, Y., & Lapata, M. (2021, August). Generating Query Focused Summaries from Query-Free Resources. In C. Zong, F. Xia, W. Li, & R. Navigli (Eds.), *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 6096–6109). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-long.475

Xu, Y., & Lapata, M. (2022). Document Summarization with Latent Queries. *Transactions of the Association for Computational Linguistics*, *10*, 623–638. https://doi.org/10.1162/tacl_a 00480

Zaheer, M., Guruganesh, G., Dubey, K. A., Ainslie, J., Alberti, C., Ontanon, S., Pham, P., Ravula, A., Wang, Q., Yang, L., & Ahmed, A. (2020). Big Bird: Transformers for Longer Sequences. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Advances in Neural Information Processing Systems* (pp. 17283–17297, Vol. 33). Curran Associates, Inc.

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2020). Bertscore: Evaluating text generation with bert. *International Conference on Learning Representations*.

Zhang, Y., Ni, A., Mao, Z., Wu, C. H., Zhu, C., Deb, B., Awadallah, A., Radev, D., & Zhang, R. (2022, May). Summˆ N: A Multi-Stage Summarization Framework for Long Input Dialogues and Documents. In S. Muresan, P. Nakov, & A. Villavicencio (Eds.), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1592–1604). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.112

Zhang, Y., Ni, A., Yu, T., Zhang, R., Zhu, C., Deb, B., Celikyilmaz, A., Awadallah, A. H., & Radev, D. (2021, November). An Exploratory Study on Long Dialogue Summarization: What Works and What's Next. In M.-F. Moens, X. Huang, L. Specia, & S. W.-t. Yih (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2021* (pp. 4426–4433). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.findings-emnlp.377

Zhu, C., Xu, R., Zeng, M., & Huang, X. (2020, November). A Hierarchical Network for Abstractive Meeting Summarization with Cross-Domain Pretraining. In T. Cohn, Y. He, & Y. Liu (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 194–203). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.findings-emnlp.19

DATASET DETAILS

Table 4. Details about the CrisisFACTS datasets. Events with the ID's 001 to 008 and 009 to 018 correspond to CrisisFACTS 2022 and CrisisFACTS 2023, respectively. The columns Tweets, Reddit, News, and Facebook represent the number of available stream items.

ID	Event	Event Type	Days	Tweets	Reddit	News	Facebook
001	Lilac Wildfire 2017	Wildfire	9	41,346	1,738	2,494	5,437
002	Cranston Wildfire 2018	Wildfire	6	22,974	231	1,967	5,386
003	Holy Wildfire 2018	Wildfire	7	23,528	459	1,495	7,016
004	Hurricane Florence 2018	Hurricane	15	41,187	120,776	18,323	196,281
005	Maryland Flood 2018	Flood	4	33,584	2,006	2,008	4,148
006	Saddleridge Wildfire 2019	Wildfire	4	31,969	244	2,267	3,869
007	Hurricane Laura 2020	Hurricane	2	36,120	10,035	6,406	9,048
008	Hurricane Sally 2020	Hurricane	8	40,695	11,825	15,112	48,492
009	Beirut Explosion 2020	Accident	7	94,892	3,257	1,163	368,866
010	Houston Explosion 2020	Accident	7	58,370	5,704	2,175	6,281
011	Rutherford TN Floods 2020	Floods	5	11,019	475	268	9,116
012	TN Derecho 2020	Storm/Flood	7	49,247	1,496	15,425	13,521
013	Edenville Dam Fail 2020	Accident	7	16,527	2,339	961	8,358
014	Hurricane Dorian 2019	Hurricane	7	86,915	91,173	7,507	370,644
015	Kincade Wildfire 2019	Wildfire	7	91,548	10,174	339	35,011
016	Easter Tornado Outbreak 2020	Tornadoes	5	91,812	5,070	750	34,343
017	Tornado Outbreak April 2020	Tornadoes	6	99,575	1,233	217	19,878
018	Tornado Outbreak March 2020	Tornadoes	6	95,221	16,911	641	87,242

PROMPT DETAILS

We provide an overview of the prompts used in our work. For the LLMNug-based methods, we rely on one demonstration example as we otherwise noticed format errors that made it difficult to parse the facts. All other methods use zero-shot prompting without the use of any demonstration examples. We list all LLM-based methods and show the prompt dependencies in the corresponding brackets: Sum (Table 5), Select+Sum (Table 5), SumCoT (Table 6, Table 7), Incremental (Table 5, Table 8), Hierarchical (Table 5, Table 9), LLMNug-Concat (Table 10), and LLMNug-ClusterFuse (Table 10, Table 5).

Table 5. Generate initial summary

We are creating one detailed summary for disaster response officers. Summarize and rewrite the following documents into coherent and readable paragraphs. Do not deviate from the facts of these documents or add any new information. Include as many details as possible. The summary could include multiple paragraphs.

Documents: {documents}

Summary:

Table 6. Element extraction

Documents: {documents}

What are the important entities in this documents?
What events are happening in this documents?
What is the result of these events?

Please answer the above questions:

Table 7. Generate element aware summary

We are creating one detailed summary for disaster response officers. Summarize and rewrite the following documents into coherent and readable paragraphs. Do not deviate from the facts of these documents or add any new information. Include as many details as possible. The summary could include multiple paragraphs. Documents: {documents} {elements} Let's integrate the above information and summarize the documents:

Table 8. Iterative refinement

We are creating one detailed summary for disaster response officers and provide an existing summary up to a certain point: {previous summary} New documents: {documents} Given the new documents, incorporate any new vital information into a new coherent and readable summary. Do not deviate from the facts of these documents or add any new information. Include as many details as possible. The summary could include multiple paragraphs. Summary:

Table 9. Hierarchical merging

We are creating one detailed summary for disaster response officers. Distill and rewrite the following previous summaries into consolidated paragraphs. Do not deviate from the facts of these summaries or add any new information. Include as many details as possible. The final summary could include multiple paragraphs. Previous summaries: {summaries} Summary:

Table 10. Fact extraction

You are a fact extractor for disaster response organizations. Use the documents to answer the question based on a list of extracted facts as evidence. Please follow the instructions for the facts: 1. The facts must be short. 2. The format of one fact is text-snippet (source document). 3. Provide the source documents for each fact with the format: (Doc-1, Doc-2, ..) 4. Include fact-relevant entities such as locations, numbers, dates, etc. 5. Only include facts which are focused on the question. 6. The list items must start with * bullet points. Do not use numberings. We provide you one example within " marks: '{demonstration}' Your task: Documents: {documents} Question: {query} Facts list:

EXAMPLES

Table 11. Summary of LLMNug-Concat w/ QILP with L=16 for request CrisisFACTS-015-r4

- Pharmacies, supermarkets, and restaurants are closed during blackouts in bedroom communities of San Francisco and Silicon Valley; Town Hall is closed except for essential services - - The main evacuation order encompasses a huge area of Sonoma County, including Santa Rosa - - Strong winds are forecasted to impact the Kincade Fire, Santa Rosa and much of the North Bay .; WILDFIRE SMOKE is impacting the San Francisco Bay Area and Los Angeles Area . - - The Kincade Fire has burned through 50,000 acres (20,200 hectares) of land; The Kincade fire spread across more than 75,000 acres (303 square kilometers) - - 4,548 firefighters are assigned to the fires; 4,500 firefighters have been working on containment lines and structure protection - - The Kincade Fire has burned 75,415 acres and is 15% contained - - The Wine & Food Affair event in Sonoma County may have been canceled; Classes at Chalon and Doheny have been canceled - - La Tortilla Factory's main office and production center is closed due to power outages and evacuations; Think Tank Photo offices are closed due to wildfires and evacuations - - At least eight homes were destroyed and six damaged in the LA-area blaze; Three people were killed and over 1,600 homes and other buildings were destroyed from north of Los Angeles to Malibu - - Four shelters for Kincade Fire evacuees are at capacity, while five are still accepting evacuees; Jaxon Keys' Winery & Distillery is an open shelter for large animals affected by the Kincade Fire - - A High Wind Warning is also in effect during the same period, with northeast winds 30 to 50 mph and gusts up to 60 to 70 mph, and isolated gusts up to 80 mph at higher elevations - - Volunteers are needed for Petaluma evacuation centers; Volunteers are needed for food service and general labor; Farm animals need volunteers to take them in - - Nearly 200,000 people have been evacuated from their homes throughout California; Over 100,000 structures are threatened - - Several state parks have closed due to the Kincade Fire and Public Safety Power Shutoffs; The registrar's office in Santa Rosa is closed because of Kincade Fire evacuations - - Red Flag No Parking restrictions are in effect citywide in Los Angeles due to the extreme fire danger - - An evacuation warning was announced for parts of Lake County; The entire town of Geyserville and the entire city of [redacted] are under mandatory evacuation orders -

Table 12. Summary of LLMNug-ClusterFuse w/ QILP with L=16 for request CrisisFACTS-015-r4

- The affected areas are located south of Mulholland, west of the 405 freeway, north of Sunset Boulevard, and west of Temescal Canyon. Entrance to Brentwood above Sunset Boulevard has been closed. - - Evacuation orders and warnings continue to apply to various cities and communities within the affected region. - - The towns of Geyserville and a redacted city, as well as certain communities and cities in the affected areas, are under mandatory evacuation orders due to ongoing wind events. - - Among them, around 4,500 firefighters have been actively engaged in setting up containment lines and safeguarding structures to prevent further damage. - - These evacuations have led to road closures on the 405 freeway, which is the nation's busiest highway. - - Nearly 1 million people are without power, and over 200,000 have been displaced from their homes. The Memorial bridge near Kennedy Ln was not impacted by the fires. - - Evacuations can pose challenges, especially for vulnerable communities that depend on childcare. Emergency response teams are encountering difficulties accessing some regions due to road closures. - - An evacuation warning has been announced for various areas in Lake County, including the northern part of Dry Creek Valley and Middletown, as well as the Twin Pines Casino area (Zone 4). - - Zone 4B, which includes areas south of Westside Road to Millcreek Road, has specific evacuation orders in effect. Some individuals have been unable to take their animals with them during evacuations. - - The registrar's office in Santa Rosa has closed due to evacuations. Recovery efforts are underway, with at least 124 structures confirmed destroyed in Sonoma County. - - Over 185,000 people have been evacuated from their homes in Sonoma County alone due to the Kincade Fire. Nearly 200,000 people have been displaced by the wildfires across California. - - Overflowing conditions have been reported in both emergency shelters and hotels, creating a critical need for additional assistance. - - Some evacuation orders have been lifted in other areas. Individuals in high-risk zones should be prepared to evacuate with short notice. - - Four of these shelters have reached full capacity, but fifteen continue to accept evacuees. Additionally, the Westwood Recreation Center serves as a shelter for those displaced by the Getty Fire. - - Sonoma County Fairgrounds and Napa Valley Expo are accommodating evacuees and small animals. Animal shelters are working diligently to provide care for pets displaced by the fires. - - Specifically mentioned are the need for volunteers at Petaluma evacuation centers, Sonoma County Animal Response Team (CART), Integrative Healers Action Network, and Kincade Fire Clinics. -

Table 13. Summary of QILP with L=16 for request CrisisFACTS-015-r4

- Some 180,000 people have been ordered to leave homes, with roads around Santa Rosa north of San Francisco packed with cars as people tried to flee. - - Evacuations have been lifted in a number of communities, and the Napa County emergency evacuation shelter at Napa Valley College closed today at noon. - - Over 200,000 people have been ordered to evacuate and more than 100,000 structures are threatened. . - - About 156,000 people were under mandatory evacuation orders. - - The following DOR offices remain closed due to the PSPS event: Novato, Ukiah and Lakeport. - - Nearly 200,000 people evacuated their homes in California this weekend. - - About 156,000 people were still under evacuation orders. - - Another million people have been told they could lose supplies. - - Evacuation Orders effecting 189,980 people. - - CWS Kits will be made available if requested, and we are checking in with our long-term recovery group partners in Los Angeles as to other ways that we can support them. - - The following road closures are in effect: Westside Road at Highway 101 offramp – no eastbound traffic on Mill Street Dry Creek Road at Highway 101 – no eastbound traffic on Dry Creek Road Chiquita Ro - - At least 10,000 people have been evacuated, gardeners and cleaners could be seen on their way to work. - - "There now are 4,548 people, mainly firefighters, assigned to the fire and more were expected later in the day." - - Areas North of Westside Road to Millcreek Road within Zone 4 have been reduced to an EVACUATION WARNING and are open for repopulation. - - Mandatory evacuations are in place for parts of the area. - - #KincadeFire VOLUNTEERS NEEDED #Petaluma People Services is coordinating volunteers for the Petaluma evacuation centers. -

Can User Characteristics Predict Norm Adherence on Social Media? Exploring User-Centric Misinformation Interventions

Shangde Gao
Department of Urban and Regional Planning
Florida Institute for Built Environment Resilience
University of Florida
gao.shangde@ufl.edu

Yan Wang
Department of Urban and Regional Planning
Florida Institute for Built Environment Resilience
University of Florida
yanw@ufl.edu

ABSTRACT

The work-in-progress explores user-centric misinformation interventions on social media as such tools are limited. Knowledge of relationships between online user characteristics and their expressed adherence to a desired norm (i.e., rejecting misinformation or supporting factual information) is understudied with limited integrated multi-modal machine-learning models to infer demographic and sociopsychological characteristics. Thus, we piloted 9,331 Twitter users tweeting COVID-19 vaccines between May 1, 2020, and April 30, 2021. Employing a CNN-LSTM framework, our model analyzes user biographies, profile images, and pre-COVID historical tweets to infer user traits over 90% accuracy for individual characteristics and an overall accuracy of 85.61%, which outperforms existing tools and other designs. Further, using multi-logistic regression, we identified significant predictors of users' adherence to desired norms, such as gender and pre-pandemic prosocial content engagement, while finding no significant age correlation. Our findings illuminate pathways for targeted misinformation mitigation strategies during critical public health crises.

Keywords

Misinformation, deep learning, social media, social norm, user-centric

INTRODUCTION

Human life has become increasingly digitized, and a massive amount of information is disseminated in an online environment. At the same time, with inaccurate and misleading content, misinformation spreads widely in the physical and digital realm and becomes a threat to society (Del Vicario et al., 2016). Misinformation can mislead individuals' perceptions and behaviors widely (Swire-Thompson & Lazer, 2020). To tackle the threat of misinformation, prebunking strategies have been proposed to achieve pre-exposure implementation and long-lasting effects on preventing the adoption of misinformation (Lewandowsky & van der Linden, 2021). Grounded in "inoculation", i.e., pre-warnings can make individuals immune from misinformation, prebunking strategies can be promising for protecting individuals from misinformation with broad and diverse topics (Cook et al., 2017; Ecker et al., 2022). Prebunking can be conducted by social norm interventions, which are particularly for population-level inoculation (Maertens et al., 2020). Social norm interventions are based on the social psychological phenomenon that people tend to conform to surrounding norms to avoid sanction or confer benefit (Constantino et al., 2022; Gao et al., 2022).

Existing practices of social norm interventions are usually designed for the general public, neglecting the heterogeneity of online users regarding their sociodemographic characteristics, psychological traits, perceptions of online information, and the capacity for detecting misinformation (Chadwick et al., 2023). For instance, questionaries among university students in Singapore have identified a significant gender difference in the

prevalence of misinformation sharing (Chen et al., 2015). A survey with 113 randomly selected social media users indicates that users between 18 and 34 years old are more likely to share misinformation (Adaji, 2023). Individuals with different influences in the online environment may also share misinformation with different tendencies (Legros & Cislaghi, 2020). Also, tendencies to generate prosocial content have been identified as influencing individuals' motivation to circulate only truthful, beneficial information (Zhu et al., 2023). Without being tailored for online users, social norm interventions may only be effective for users with certain characteristics instead of the whole population of online users.

Compared to interventions that are generally designed for the public, user-centric or tailored interventions are designed to align with the heterogeneous characteristics of online users and may potentially amplify the effectiveness of interventions (Kaufman et al., 2022). User-centric or tailored designs of interventions have been successfully applied in other domains of information dissemination and communications, such as tailored and targeted communications of hazard information and health-related information (Gao & Wang, 2021; Kreuter & Wray, 2003). However, existing models that estimate social media user characteristics have not reached high-level accuracy (Wang et al., 2019). They also have not been designed to estimate users' social psychological characteristics, such as tendencies to generate prosocial content. Also, existing studies have not comprehensively investigated the relationship between user characteristics and behaviors of sharing or rejecting misinformation, providing insufficient knowledge about selecting targeted users who may need to be motivated to adhere to the *desired norm*, expressed in tweets of rejecting online misinformation or supporting the factual information.

We aim to answer the following two research questions:

RQ 1: How can we accurately infer online users' demographic and social psychological characteristics based on their features in profile images, biography text, and online posts?

RQ 2: What are the relationships between online user characteristics (including age, gender, tendency to generate prosocial content, and influence on social media) and users' adherence to the desired norms?

Focusing on the topic of COVID-19 vaccines on Twitter, we start by capturing 9,331 sample users' expressed adherence to the desired norm from May 1, 2020, to April 30, 2021. We infer online users' demographic and social psychological characteristics by developing a multi-modal deep learning model. The input data includes users' profile information and pre-COVID tweets posted between January 1, 2019, and December 31, 2019. Then, we use multi-logistic regression to examine the relationship between user characteristics and adherence to the desired norm. The preliminary study paves the way for the design of user-centric social norm interventions and contributes to the existing knowledge body in crisis informatics in two ways. First, it develops a comprehensive model designed to accurately infer a diverse set of user characteristics by analyzing their online profiles and social media posts. By integrating different data features, the model processes textual, visual, and behavioral data to predict user traits with high accuracy. Second, our analysis reveals significant correlations between user characteristics and their propensity to adhere to the desirable norm. The findings offer actionable insights for developing user-centric misinformation mitigation strategies, particularly critical in preemptive responses to future pandemics and large-scale crises.

LITERATURE REVIEW

Existing Prediction Models of Online User Characteristics

Due to the privacy policies of online platforms, the demographic characteristics of online users are usually not recorded and are not publicly accessible (Wang et al., 2019). Existing studies have proposed various methods for predicting online user characteristics, mainly focusing on demographic characteristics, especially genders and ages (or age groups). The predictions are mainly based on biographies, online posts, or profile images of online users, and the basic method of prediction is machine learning (especially deep learning) models. For example, support vector machine (SVM) has been widely used to predict the gender and age groups of users on Twitter and Facebook (Chen et al., 2021; Fink et al., 2021). Deep learning models have also been proposed to infer the gender and age of online users based on users' profile information, such as the M3 inference, which is a multi-modal deep learning model utilizing users' profile information (Wang et al., 2019).

However, existing tools may not be effective at predicting the characteristics of English-speaking users due to the biases in input features and datasets. For example, some tools rely on profile images as the primary feature for inferring users' demographics (Golder et al., 2022; Priadana et al., 2020), and they may not perform well with users with default profile images or no profile images. Also, some tools are developed based on certain platforms, such as Instagram and Facebook (Çoban et al., 2021; Priadana et al., 2020), which may affect their accuracy in inferring the characteristics of other platforms, such as Twitter. In another case, as one of the models with the

highest inference accuracy, M3 inference can still be improved to predict English Twitter users' age and gender, as its current accuracy for age prediction is around 40% and for gender is around 80%.

Meanwhile, online users' social psychological behaviors, such as tendencies to generate prosocial content, have not been predicted by previous studies. In particular, previous studies have proposed questionnaires to collect individuals' responses to some sociopsychological questions and estimated their tendencies to engage in prosocial behaviors accordingly (Lavertu et al., 2020; Perach et al., 2023). This strategy is not suitable for large populations due to the cost and may not apply to online users who are anonymous and may not participate in online surveys.

Online User Characteristics and Information Dissemination

Online users' demographic characteristics have been found to be related to their behaviors of information dissemination and the spreading of misinformation in the online environment. In some experiments with randomly sampled social media users, it was observed that female users and users between the ages of 18 and 34 have shown relatively higher tendencies to share misinformation without detecting it (Balakrishnan, 2022; Xiang et al., 2023). In addition to demographic characteristics, individuals' perceptions and dissemination of misinformation may be influenced by their social psychological traits, such as the fear of being attacked, desire to protect their self-image, lack of self-efficacy, lack of accountability, personal traits, extraversion, empathic concern, and tendencies to being prosocial (Gurgun et al., 2022; Perach et al., 2023). Among these characteristics, the tendency to be prosocial may be related to detecting misinformation, according to the experiments in China and Hungary (Orosz et al., 2023; Sun & Ma, 2023; Zhu et al., 2023). Behaviors of being prosocial include but are not limited to generating content for supporting or helping others in pursuit of developing and maintaining harmonious relationships (Van Rijsewijk et al., 2016). Online users with varying levels of influence in online environments may also have different tendencies to share or reject misinformation. Some online users, such as online news media, celebrities, journalists, and other populations with large followings, may have a greater impact on online discussions than the ones with low-level influence (Andrighetto et al., 2013; Legros & Cislaghi, 2020). Low-influence users' attitudes toward online misinformation are likely to align with the norms that are proposed and shaped by high-influence users (Constantino et al., 2022; Wu et al., 2011).

However, studies on the relationships between user characteristics and the adoption of misinformation-related beliefs may still be limited due to their reliance on self-reported data from survey participants, while individuals' responses in the surveys may deviate from their actual behaviors and beliefs (Takeuchi et al., 2023). Online users' historical posts can provide observed information about users' tendency to reject or share misinformation, but such data has not been sufficiently investigated by existing studies. Also, existing findings about the relationships between being prosocial and detecting misinformation may not be fully applied to English-speaking Twitter users, who are not covered by their experiment samples. Additionally, existing studies have mainly examined users' tendency or behaviors to detect misinformation, which does not always equal to rejecting misinformation content or expressing adherence to the norm of rejecting misinformation or supporting factual information (Moravec et al., 2018).

METHODOLOGY

Data Collection and Preprocessing

We collected Twitter data related to COVID-19 using keywords such as "covid", "coronavirus", and "sars-cov-2" through Twitter's publicly available Streaming API, as detailed in our previous studies (Gao et al., 2022; Gao et al., 2023; Wang et al., 2017; Wang et al., 2021). To ensure privacy, results were reported only at the population level, addressing potential data privacy concerns. From the large dataset, we randomly sampled 10,000 users who tweeted about COVID-19 vaccines between May 1, 2020, and April 30, 2021. After removing suspended or protected user accounts, 9,331 users remained in our study sample. The data set includes user profiles, especially profile images and biographies. Additionally, we retrieved the sample users' historic tweets from the pre-COVID period between January 1, 2019, and December 31, 2019, using a web-scraping tool named Nitter Scraper (DGNSREKT, 2023). Both the user profile and historical tweets were employed to infer demographic characteristics and tendencies of generating prosocial content using a CNN-LSTM-based multi-modal model. Notably, we leveraged a different dataset – Wikidata – to train the model because the dataset includes self-reported year-of-birth and genders of 6,739 users (Liu & Singh, 2021). Details can be found in the Model training section. Though some users may have deleted their posts, this retrospective data collection approach relies on the assumption that tendencies of being prosocial are consistent over time (Eisenberg, 2023), minimizing the impact of deleted posts on the overall analysis.

Developing an Integrated Model to Uncover Users' Demographic and Social-Psychological Characteristics

We developed an integrated multi-modal model (Figure. 1) to infer online users' gender (male or female), age groups (under 18, 18 to 44, 45 to 64, 65 or older), and tendency to generate prosocial content ("yes" or "no").

Model Input

The input of our integrated model includes the profile images and users' biographies and pre-COVID tweets, i.e., the text, topics, and sentiment, which can be distinguished among users with different ages and genders (Holmberg & Hellsten, 2015; Nguyen et al., 2021). The user images were directly input to the image-related module. The text of users' biographies and pre-COVID tweets were preprocessed before inputting into the text-related module by tokenizing words, lowering each word, and removing English stopwords, punctuations, URLs, and @usernames (Yao & Wang, 2020). After preprocessing, the text of biographies and pre-COVID tweets were embedded into two vectors based on the corpus of biographies and pre-COVID tweets of all users, then input into the text-related module. Specifically, the embedded process converted each piece of input text into a dense vector, in which each element represented the index of a word in the corpus with a 1×60 fixed-size embedding. The topics of users' biographies and the topics of their pre-COVID tweets were determined by Latent Dirichlet Allocation (LDA) topic modeling for each user (Yao & Wang, 2020). We set the possible count of topics as eight to obtain the lowest perplexity score of topic modeling. The LDA algorithm started by calculating the probability of each word in the corpus belonging to a certain topic and assigning the words to the topics with the highest probability. Each user's biography and generated tweets were then assigned the topics with the highest frequency in their respective texts. The sentiments of users' biographies and pre-COVID tweets were detected with VADER, a keyword-based and rule-based model (Hutto & Gilbert, 2014). The topics and sentiments of users' biographies and pre-COVID tweets were also input into the text-related module of our model.

Model Structure

The multi-modal model comprises two pipelines, which process the image input and the text-based inputs separately. The first pipeline includes a Lenet-based convolutional neural network (CNN) module, which is widely used for processing imagery inputs (Lecun et al., 1998). The CNN module consists of three sets of "convolution + max pooling" layers for processing the raw profile images, of which the size is 400×400. These layers are followed by a flatting layer that turns the 2D data into a 1D vector, which is the output of the first pipeline.

The second pipeline contains two parallel processes for biography-related inputs and pre-COVID tweet-related inputs. For the biography-related inputs, the word vector is processed by a Long-Short Term Memory (LSTM) layer, which is typically utilized for capturing information from text data (Sundermeyer et al., 2012). The topic and sentiment of user biographies are transformed into two dummy vectors. The output vector of the LSTM layer and the dummy vectors of topics and sentiments are then aggregated with a dense layer, which reduces the total dimension of these three vectors and avoids co-linearity. The processing of pre-COVID tweet-related inputs is the same as the processing of biography-related inputs. The outputs of the second pipeline include one vector representing users' biographies and one vector representing users' pre-COVID tweets.

The output vectors of the two pipelines are concatenated into one vector and input into a three-layer feedforward neural network, which generates a dummy variable indicating the user's characteristics. By considering users' gender (two types), age groups (four types), and tendency to generate prosocial content (two types), an online user can belong to one of 16 classes ($2 \times 4 \times 2$). For example, one class of users can be male, under 18, and tend to generate prosocial content.

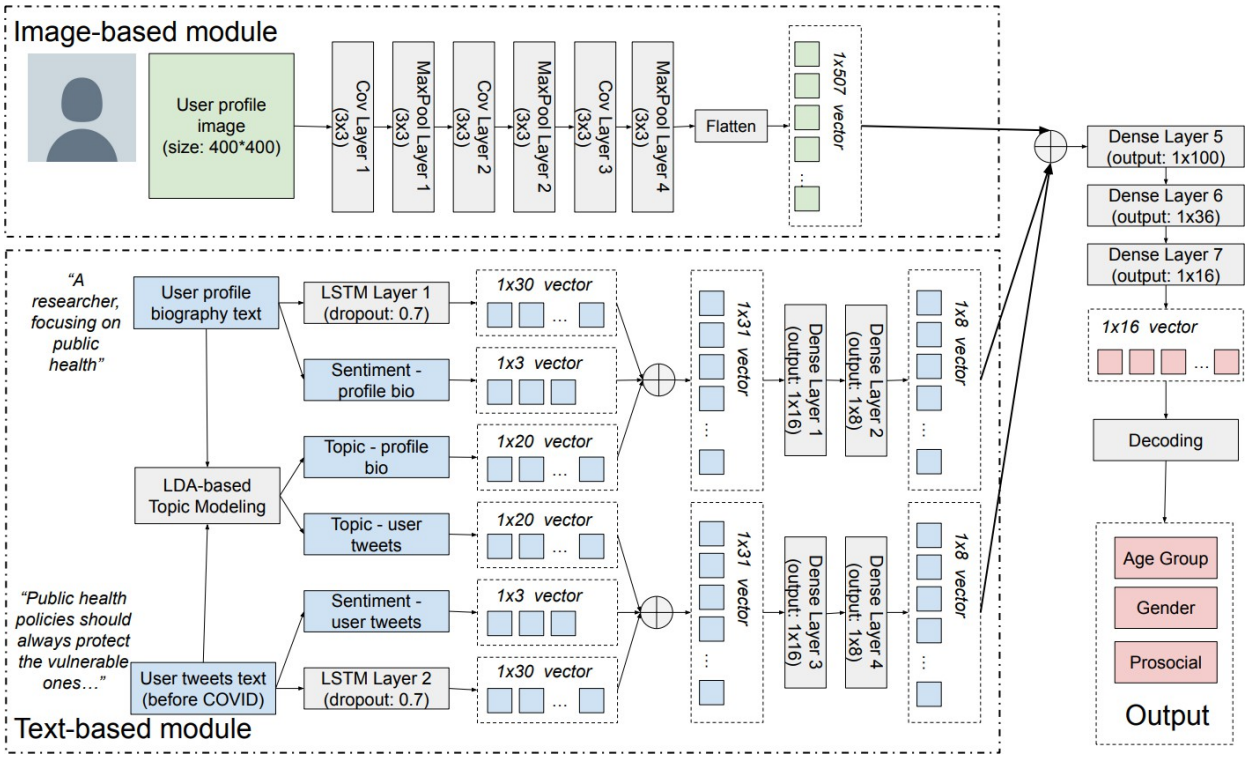


Figure 1. Architecture of the multi-modal model

Model Training

To train the model shown in Figure. 1, we used Wikidata as the training dataset containing users with self-reported year-of-birth and gender (Liu & Singh, 2021). We utilized the Nitter scraper to retrieve profile images, biographies, and pre-COVID tweets generated by the users in Wikidata between January 1, 2019, and December 1, 2019. We can retrieve the complete information from 6,739 users from Wikidata. To determine whether the combined corpora of the tweets and biographies of users in training data were prosocial or not, we labeled each user's tweet corpora based on manual reviewing and the judgment of GPT-4 (OpenAI, 2023), following existing studies (Li et al., 2024; Mujahid et al., 2023). A text was labeled as prosocial content if it contained positive, friendly, supporting, and gratitude expressions (Lysenstøen et al., 2021). An example is "How great would it be if we took the time to sit down with our veterans, visit and thank them for their service." The distribution of users in each class in the training data is shown in Figure. 2.

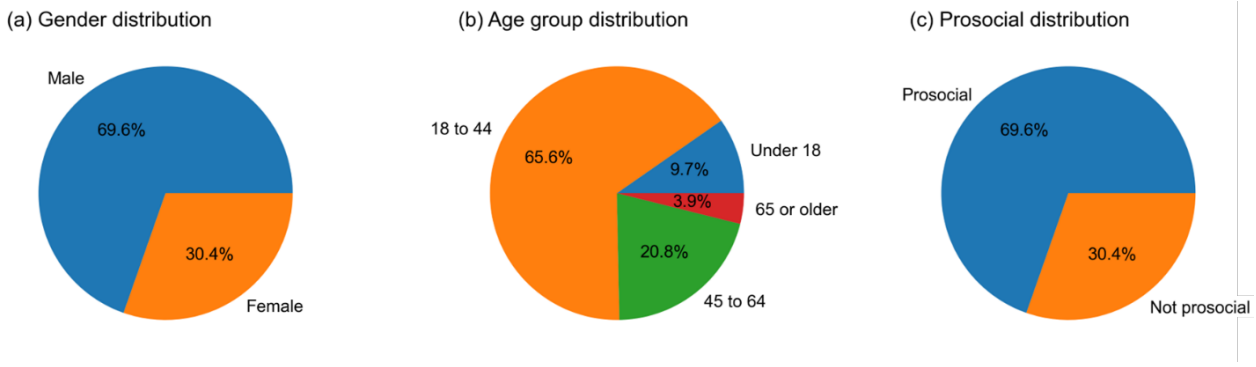


Figure 2. Distribution of genders, age groups, and users generating prosocial content in the training data

To train the multi-modal model, we split the training data into two sets, which separately took 70% (for modeling training) and 30% (for model validation) of the training dataset. The training process consisted of 20 epochs to avoid overfitting. The batch size was set to 100 to reach the balance between the times of updating model parameters and model accuracy (Smith et al., 2017). We selected Adam as the optimizer and calculated the model loss with categorical cross-entropy.

Tracking Online Users' Expressed Adherence to the Desired Norm

Leveraging the trained multi-modal model in the prior session, we can predict the user characteristics of 9,331 users. Then, we determined each user's expressed adherence to the desired norm. We started by determining if each tweet of each user expressed adherence to the desired norm. A tweet was considered as not adhering to the desired norm if its content delivered manipulated information about the side effects of COVID-19 vaccines, such as "do not take COVID-19 vaccines, they made children killed." In contrast, we considered a tweet as adhering to the desired norm if it supported COVID-19 vaccination or rejected related misinformation, such as "take COVID-19 vaccines to protect you and your family." We trained an LSTM classifier for the text classification based on 2,000 manually labeled tweets (Table 1) (Gao et al., 2023). The tweets were randomly sampled from all the real-time tweets we previously collected, which were not generated by our studied 9,331 users. LSTM was used because of its high capacity to capture phrase-level and sentence-level feature patterns (Sundermeyer et al., 2012). The validated accuracy and loss of the LSTM classifier during training are shown in Figure. 3, reaching 0.8892 and 0.2292 separately after training. Also, the RMSE of the classification outcomes is 0.3719, and the F1 score is 0.8706. These metrics indicate that our LSTM classifier has a satisfactory performance.

Table 1. Proportions of tweets expressing or not expressing adherence to the desired norm in training and testing data of the LSTM classifier

Dataset	Size	Count/Proportion of Tweets Expressing Adherence to the Desired Norm	Count/Proportion of Tweets Not Expressing Adherence to the Desired Norm
Training data	1,600 tweets	209 tweets (13.06%)	1,391 tweets (86.94%)
Testing data	400 tweets	57 tweets (14.25%)	343 tweets (85.75%)

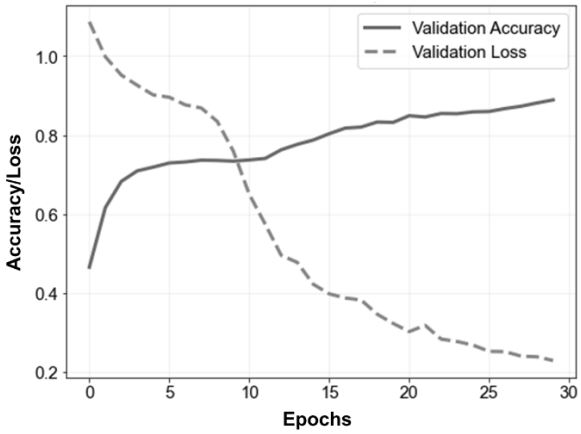


Figure 3. The accuracy and loss of the LSTM classifier during training

After determining users' expressed beliefs delivered in each tweet, we specified each user's expressed adherence to the desired norm in the twelve-month studied period (represented by $AAAAheeeeeeeeeeee$). We first specified their vaccination-relevant tweets each month. We labeled users' tweets corpora in one month as expressing adherence to the desired norm if all of their tweets about COVID-19 vaccines in that month expressed adherence to the desired norm. Based on the monthly adherence to the desired norm, we categorized users into two groups: users consistently expressing adherence to the desired norm ($AAAAheeeeeeeeeeee = 1$), and users not consistently expressing adherence to the desired norm ($AAAAheeeeeeeeeeee = 0$).

Multi-Logistic Regression for Examining the Relationship between Online User Characteristics and Expressed Adherence to the Desired Norm

We examined the relationship between the user characteristics (age, gender, tendency to generate prosocial content, and count of followers) and $AAAAheeeeeeeeeeee$ with multi-logistic regression based on 9,331 Twitter users (**Eq. 1**). One characteristic of online users is considered as significantly correlated with $AAAAheeeeeeeeeeee$ if the $pp - vvvvvvvvee$ of its coefficient is lower than 0.01 or 0.001, representing low to high levels of significance separately.

$$AAAAheeeeeeeeeeee \sim AAAAee + GGeeeeAAeeee + PPeePPPPPPeePPvvvv + IIeeIIvvvveeeeeeee \quad (1)$$

RESULTS

Prediction Outcomes of User Characteristics and Monthly Patterns of Expressed Adherence to the Desired Norm

With the multi-modal model, we inferred users' genders, age groups, and tendencies to generate prosocial content. The statistical distribution of user characteristics is shown in Figure. 4, which is similar to the distribution among the users in the training data. 65.0% of the analyzed users are male. Most of the analyzed users belong to the age group of 18-44 (77.0%) or 45-64 (17.2%). 77.5% of the analyzed users tend to generate prosocial content.

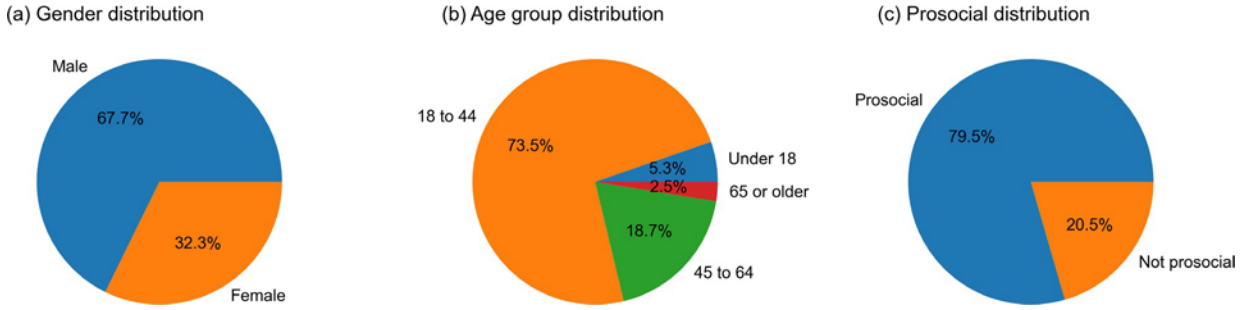


Figure 4. Distribution of genders, age groups, and users generating prosocial content among the analyzed users

After estimating the analyzed users' expressed adherence to the desired norm, we categorized users into two groups (Figure. 5a). While most of our analyzed users posted less than 200 tweets in the studied period (Figure. 5b), 44.3% of the analyzed users consistently expressed adherence to the desired norm.

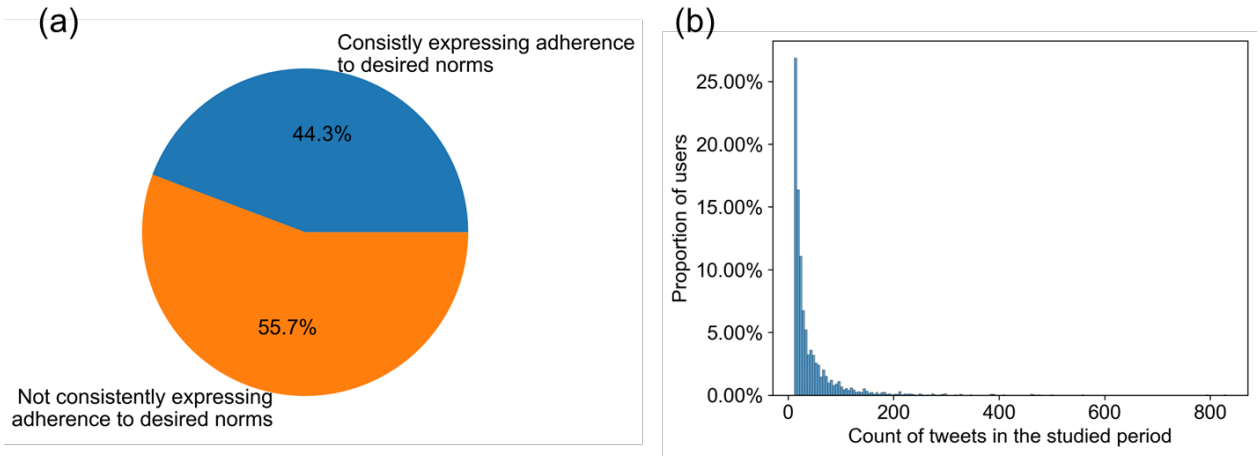


Figure 5. Proportions of users who consistently expressed adherence to the desired norm (a) and proportion of users with different counts of tweets

Evaluation of Multi-Modal CNN-LSTM Model Performance

We evaluated the multi-modal model that infers online user characteristics based on its accuracy and loss and ablation studies. In terms of accuracy and loss, our model achieved an accuracy of 85.61% when predicting online users' gender, age group, and tendency to generate prosocial content simultaneously, and its accuracy exceeds 90% for predicting each characteristic separately (Figure. 6 and Table 2). We also conducted ablation studies (Table 2), i.e., examining the prediction performance of each component of the multi-modal model and their combinations, to investigate their contributions to the prediction process. The multi-modal model has a much higher accuracy than each component classifier and each of their combinations. Among the model components, the image-based module (CNN classifier) has the major contribution and can solely achieve an overall accuracy of 60.83%, which is much higher than the text-based module.

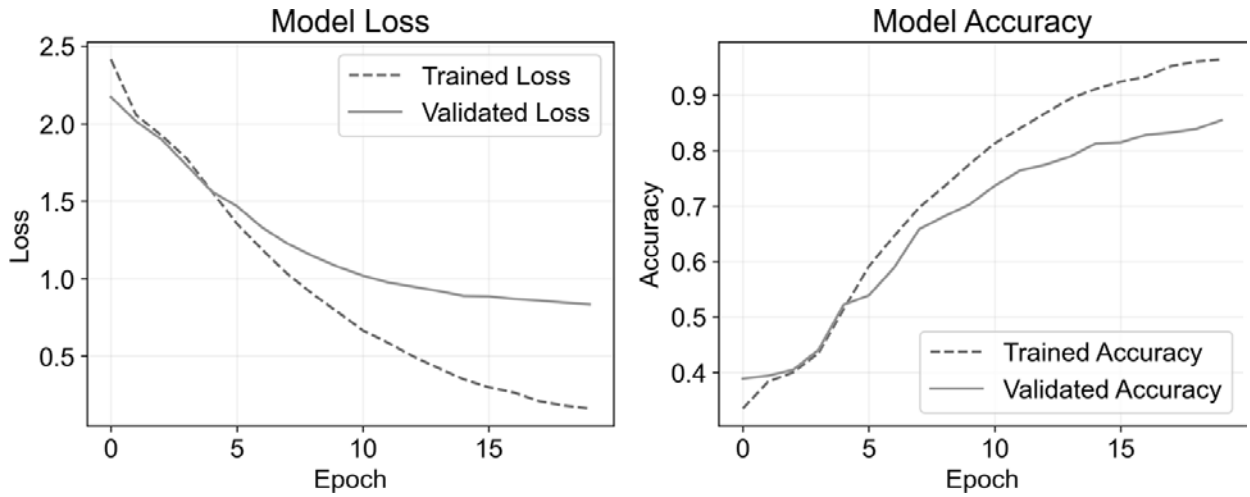


Figure 6. Loss and accuracy of the multi-modal model

Table 2. Ablation studies of the multi-modal model

Model	Inferring gender	Inferring age groups	Inferring prosocial	Overall
CNN (profile image)	0.7597	0.7538	0.8426	0.6083
LSTM (profile bio)	0.7284	0.6683	0.8029	0.5194
LSTM (tweet text)	0.6889	0.6577	0.8012	0.4486
CNN (profile image) + LSTM (profile bio)	0.7717	0.7200	0.8139	0.5710
CNN (profile image) + LSTM (tweet text)	0.7581	0.7084	0.7993	0.5480
LSTM (profile bio) + LSTM (tweet text)	0.7421	0.6889	0.8059	0.4753
Proposed model	**0.9069**	**0.9162**	**0.9467**	**0.8561**

Correlation between Online User Characteristics and Expressed Adherence to the Desired Norm

Before conducting correlation analysis, we examined the inter-correlation between online user characteristics and indicators of influence in online environments (Figure. 7). Notably, the user characteristics and influence in online environments are not significantly correlated and can be input into the regression analysis.

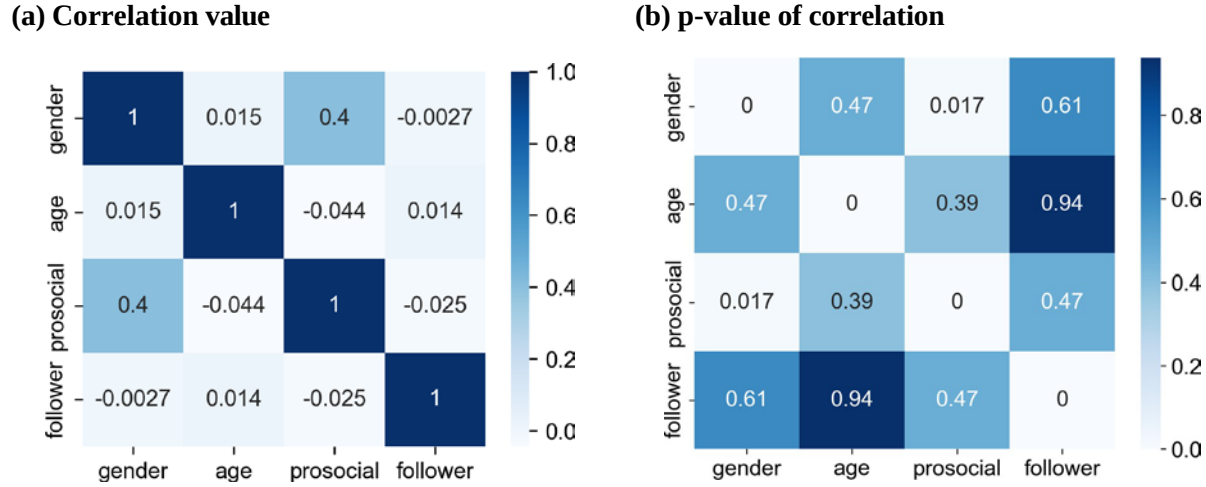


Figure 7. Correlation values and $pp - vvvvvvvvv$ between user characteristics

We then conducted multi-logistic regression to examine the relationship between user characteristics and users' adherence to the desired norm ($AAAAheeeeeeeeeeee$) (Table 3). The regression outcomes indicate that gender ($pp -$

$vvvvvvvvee = 0.008$) and the tendency to generate prosocial content ($pp - vvvvvvvvee = 0.003$) are significantly correlated with consistently expressed adherence to the desired norm. A user is more likely to consistently express adherence to the desired norm if the user is male or tends to generate prosocial content before the pandemic. In contrast, female users and users who do not generate prosocial content are less likely to express adherence to the desired norm.

Table 3. Multi-logistic regression outcomes

	Coefficient	**Confident Interval (lower 2.5% boundary)**	**Confident Interval (upper 97.5% boundary)**	$pp - vvvvvvvvv$
Intercept	-0.6695	-1.2020	-0.1370	0.0140
Gender	**-0.3097**	**-0.5380**	**-0.0820**	**0.0080****
Age	-0.1283	-0.3230	0.0660	0.1960
Prosocial	**0.3792**	**0.1320**	**0.6260**	**0.0030****
Count of Followers	0.0000	0.0000	0.0000	0.3250
Significance Levels: 0 '***' 0.001 '**' 0.01				

DISCUSSION AND FUTURE WORK

The research is among the first to explore the idea of user-centric normative misinformation intervention. The preliminary study contributes to the knowledge body of crisis informatics in online environments through two major outcomes. First, we address the need for an integrated online user characteristics inference model and propose a new multi-modal deep-learning model for predicting demographic and social psychological characteristics. With an overall accuracy of 85.61% and an accuracy exceeding 90% for each user characteristic, our model outperforms existing tools for inferring online users' demographic and social psychological characteristics based on publicly accessible data. Second, our statistical analysis has identified a significant and positive relationship between users' gender and their tendency to generate prosocial content and expressed adherence to the desired norm. The correlation highlights the necessity of paying attention to the users who are female and do not tend to generate pro-social content for social norm interventions. We expect that more customized and user-centric prebunking social norm interventions will alleviate the adverse social consequences of misinformation more effectively in the digital environment.

Our further research will continue to explore and test user-centric misinformation mitigations to advance the knowledge body of crisis informatics following the work-in-progress paper. First, we will expand the scope of online user characteristics for predicting online users' expressed adherence to the desired norm and misinformation content, such as self-esteem, optimism, and expressed satisfaction with daily life. Second, we will integrate the impact of individuals' exposure to their information environments to understand the motivation for changing expressed belief in vaccination. Particularly, online users' adoption of the desired norm is potentially related to the expressed beliefs of their social network members. Third, we will investigate other topics of online misinformation, such as healthy eating, threats and protective measures against communicable diseases, and risky behaviors (such as the use of drugs and smoking). Lastly, our future work will focus on designing user-centric interventions to mitigate online misinformation before the next pandemic or crisis. For example, we will explore the strategies to break the "echo chamber" among users who do not tend to generate prosocial content. We will also examine the effectiveness of spreading messages that deliver the desired norm or prosocial content to encourage users to reject misinformation.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 2323794. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Adaji, I. (2023). Age Differences in the Spread of Misinformation Online. *European Conference on Social Media*, 10(1), 12–19. https://doi.org/10.34190/ecsm.10.1.1156

Andrighetto, G., Castelfranchi, C., Mayor, E., McBreen, J., Lopez-Sanchez, M., & Parsons, S. (2013). (Social) Norm Dynamics. *Normative Multi-Agent Systems*, 135–170. https://doi.org/10.4230/DFU.VOL4.12111.135

Balakrishnan, V. (2022). Socio-demographic Predictors for Misinformation Sharing and Authenticating amidst the COVID-19 Pandemic among Malaysian Young Adults. *Information Development*, 026666692211189. https://doi.org/10.1177/02666669221118922

Chadwick, A., Vaccari, C., & Hall, N.-A. (2023). What Explains the Spread of Misinformation in Online Personal Messaging Networks? Exploring the Role of Conflict Avoidance. *Digital Journalism*, 1–20. https://doi.org/10.1080/21670811.2023.2206038

Chen, X., Sin, S.-C. J., Theng, Y.-L., & Lee, C. S. (2015). Why Do Social Media Users Share Misinformation? *Proceedings of the 15th ACM/IEEE-CS Joint Conference on Digital Libraries*, 111–114. https://doi.org/10.1145/2756406.2756941

Chen, X., Wang, Y., Agichtein, E., & Wang, F. (2021). A Comparative Study of Demographic Attribute Inference in Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, *9*(1), 590–593. https://doi.org/10.1609/icwsm.v9i1.14656

Çoban, Ö., İnan, A., & Özel, S. A. (2021). Facebook Tells Me Your Gender: An Exploratory Study of Gender Prediction for Turkish Facebook Users. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *20*(4), 1–38. https://doi.org/10.1145/3448253

Constantino, S. M., Sparkman, G., Kraft-Todd, G. T., Bicchieri, C., Centola, D., Shell-Duncan, B., Vogt, S., & Weber, E. U. (2022). Scaling Up Change: A Critical Review and Practical Guide to Harnessing Social Norms for Climate Action. *Psychological Science in the Public Interest*, *23*(2), 50–97. https://doi.org/10.1177/15291006221105279

Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLOS ONE*, *12*(5), e0175799. https://doi.org/10.1371/journal.pone.0175799

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

DGNSREKT. (2023). Nitter scraper. https://nitter-scraper.readthedocs.io/en/latest/#docker-engine

Ecker, U. K., Sanderson, J. A., McIlhiney, P., Rowsell, J. J., Quekett, H. L., Brown, G. D., & Lewandowsky, S. (2022). Combining refutations and social norms increases belief change. *Quarterly Journal of Experimental Psychology*, 174702182211117. https://doi.org/10.1177/17470218221111750

Eisenberg, N. (2003). Prosocial behavior, empathy, and sympathy. In *Well-Being Positive Development Across the Life Course*. Taylor and Francis.

Fink, C., Kopecky, J., & Morawski, M. (2021). Inferring Gender from the Content of Tweets: A Region Specific Example. *Proceedings of the International AAAI Conference on Web and Social Media*, *6*(1), 459–462. https://doi.org/10.1609/icwsm.v6i1.14320

Gao, S., & Wang, Y. (2021). Assessing the impact of geo-targeted warning messages on residents' evacuation decisions before a hurricane using agent-based modeling. *Natural Hazards*, *107*(1), 123–146. https://doi.org/10.1007/s11069-021-04576-1

Gao, S., Wang, Y., & Thai, M. T. (2023). Investigating the Dynamics of Social Norm Emergence within Online Communities. https://doi.org/10.48550/ARXIV.2301.00453

Gao, S., Wang, Y., & Webster, G. D. (2022). Causal Modeling of Descriptive Social Norms from Twitter and the Physical World on Expressed Attitudes Change: A Case Study of COVID-19 Vaccination. *Cyberpsychology, Behavior, and Social Networking*, *25*(12), 769–775. https://doi.org/10.1089/cyber.2022.0153

Golder, S., Stevens, R., O'Connor, K., James, R., & Gonzalez-Hernandez, G. (2022). Methods to Establish Race or Ethnicity of Twitter Users: Scoping Review. *Journal of Medical Internet Research*, *24*(4), e35788. https://doi.org/10.2196/35788

Gurgun, S., Arden-Close, E., Phalp, K., & Ali, R. (2022). Online silence: Why do people not challenge others when posting misinformation? *Internet Research*. https://doi.org/10.1108/INTR-06-2022-0407

Holmberg, K., & Hellsten, I. (2015). Gender differences in the climate change communication on Twitter. *Internet Research, 25*(5), 811–828. https://doi.org/10.1108/IntR-07-2014-0179

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Jiang, L. C., Chu, T. H., & Sun, M. (2021). Characterization of Vaccine Tweets During the Early Stage of the COVID-19 Outbreak in the United States: Topic Modeling Analysis. *JMIR Infodemiology, 1*(1), e25636. https://doi.org/10.2196/25636

Kaufman, J., Bagot, K. L., Tuckerman, J., Biezen, R., Oliver, J., Jos, C., Ong, D. S., Manski-Nankervis, J., Seale, H., Sanci, L., Munro, J., Bell, J. S., Leask, J., & Danchin, M. (2022). Qualitative exploration of intentions, concerns and information needs of vaccine-hesitant adults initially prioritised to receive COVID-19 vaccines in Australia. *Australian and New Zealand Journal of Public Health, 46*(1), 16–24. https://doi.org/10.1111/1753-6405.13184

Kreuter, M., & Wray, R. (2003). Tailored and targeted health communication: Strategies for enhancing information relevance. *American Journal of Health Behavior*, 27, S227–S232. https://doi.org/10.5993/ajhb.27.1.s3.6

Lavertu, L., Marder, B., Erz, A., & Angell, R. (2020). The extended warming effect of social media: Examining whether the cognition of online audiences offline drives prosocial behavior in 'real life.' *Computers in Human Behavior*, 110, 106389. https://doi.org/10.1016/j.chb.2020.106389

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Legros, S., & Cislaghi, B. (2020). Mapping the Social-Norms Literature: An Overview of Reviews. *Perspectives on Psychological Science, 15*(1), 62–80. https://doi.org/10.1177/1745691619866455

Lewandowsky, S., & van der Linden, S. (2021). Countering Misinformation and Fake News Through Inoculation and Prebunking. *European Review of Social Psychology, 32*(2), 348–384. https://doi.org/10.1080/10463283.2021.1876983

Li, L., Fan, L., Atreja, S., & Hemphill, L. (2024). "HOT" ChatGPT: The Promise of ChatGPT in Detecting and Discriminating Hateful, Offensive, and Toxic Comments on Social Media. *ACM Transactions on the Web*, 3643829. https://doi.org/10.1145/3643829

Liu, Y., & Singh, L. (2021). Age Inference Using A Hierarchical Attention Neural Network. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 3273–3277. https://doi.org/10.1145/3459637.3482055

Lysenstøen, C., Bøe, T., Hjetland, G. J., & Skogen, J. C. (2021). A Review of the Relationship Between Social Media Use and Online Prosocial Behavior Among Adolescents. *Frontiers in Psychology, 12*, 579347. https://doi.org/10.3389/fpsyg.2021.579347

Maertens, R., Anseel, F., & van der Linden, S. (2020). Combatting climate change misinformation: Evidence for longevity of inoculation and consensus messaging effects. *Journal of Environmental Psychology, 70*, 101455. https://doi.org/10.1016/j.jenvp.2020.101455

Minaei, M., Mouli, S. C., Mondal, M., Ribeiro, B., & Kate, A. (2021). Deceptive Deletions for Protecting Withdrawn Posts on Social Media Platforms. *Proceedings 2021 Network and Distributed System Security Symposium. Network and Distributed System Security Symposium, Virtual.* https://doi.org/10.14722/ndss.2021.23139

Moravec, P., Minas, R., & Dennis, A. R. (2018). Fake News on Social Media: People Believe What They Want to Believe When it Makes No Sense at All. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3269541

Mujahid, M., Kanwal, K., Rustam, F., Aljedaani, W., & Ashraf, I. (2023). Arabic ChatGPT Tweets Classification Using RoBERTa and BERT Ensemble Model. *ACM Transactions on Asian and Low-Resource Language Information Processing, 22*(8), 1–23. https://doi.org/10.1145/3605889

Nguyen, D., Gravel, R., Trieschnigg, D., & Meder, T. (2021). "How Old Do You Think I Am?" A Study of Language and Age in Twitter. *Proceedings of the International AAAI Conference on Web and Social Media, 7*(1), 439–448. https://doi.org/10.1609/icwsm.v7i1.14381

OpenAI. (2023). *GPT-4: Generative Pre-trained Transformer 4.* Retrieved from https://openai.com/gpt-4.

Orosz, G., Paskuj, B., Faragó, L., & Krekó, P. (2023). A prosocial fake news intervention with durable effects. *Scientific Reports, 13*(1), 3958. https://doi.org/10.1038/s41598-023-30867-7

Perach, R., Joyner, L., Husbands, D., & Buchanan, T. (2023). Why Do People Share Political Information and Misinformation Online? Developing a Bottom-Up Descriptive Framework. *Social Media + Society*, *9*(3), 20563051231192032. https://doi.org/10.1177/20563051231192032

Priadana, A., Maarif, M. R., & Habibi, M. (2020). Gender Prediction for Instagram User Profiling using Deep Learning. *2020 International Conference on Decision Aid Sciences and Application (DASA)*, 432–436. https://doi.org/10.1109/DASA51403.2020.9317143

Smith, S. L., Kindermans, P.-J., Ying, C., & Le, Q. V. (2017). Don't Decay the Learning Rate, Increase the Batch Size. https://doi.org/10.48550/ARXIV.1711.00489

Sun, M., & Ma, X. (2023). Combating health misinformation on social media through fact-checking: The effect of threat appraisal, coping appraisal, and empathy. *Telematics and Informatics*, *84*, 102031. https://doi.org/10.1016/j.tele.2023.102031

Sundermeyer, M., Schlüter, R., & Ney, H. (2012). LSTM neural networks for language modeling. *Interspeech 2012*, 194–197. https://doi.org/10.21437/Interspeech.2012-65

Swire-Thompson, B., & Lazer, D. (2020). Public Health and Online Misinformation: Challenges and Recommendations. *Annual Review of Public Health*, *41*(1), 433–451. https://doi.org/10.1146/annurev-publhealth-040119-094127

Takeuchi, M., Niimi, J., & Hoshino, T. (2023). Handling the Inconsistency between Self-Report and the Actual Behavior: Validity of Excluding Survey Participants with Insufficient Effort Responding. *International Journal of Market Research*, 14707853231209933. https://doi.org/10.1177/14707853231209933

Van Rijsewijk, L., Dijkstra, J. K., Pattiselanno, K., Steglich, C., & Veenstra, R. (2016). Who helps whom? Investigating the development of adolescent prosocial relationships. *Developmental Psychology*, *52*(6), 894–908. https://doi.org/10.1037/dev0000106

Wang, Y., Wang, Q., & Taylor, J. E. (2017). Aggregated responses of human mobility to severe winter storms: An empirical study. *PLOS ONE*, *12*(12), e0188734. https://doi.org/10.1371/journal.pone.0188734

Wang, Z., Hale, S., Adelani, D. I., Grabowicz, P., Hartman, T., Flöck, F., & Jurgens, D. (2019). Demographic Inference and Representative Population Estimates from Multilingual Social Media Data. *The World Wide Web Conference*, 2056–2067. https://doi.org/10.1145/3308558.3313684

Wang, Y., Hao, H., & Platt, L. S. (2021). Examining risk and crisis communications of government agencies and stakeholders during early-stages of COVID-19 on Twitter. *Computers in Human Behavior*, *114*, 106568. https://doi.org/10.1016/j.chb.2020.106568

Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Who says what to whom on twitter. *Proceedings of the 20th International Conference on World Wide Web*, 705. https://doi.org/10.1145/1963405.1963504

Xiang, H., Zhou, J., & Wang, Z. (2023). Reducing Younger and Older Adults' Engagement with COVID-19 Misinformation: The Effects of Accuracy Nudge and Exogenous Cues. *International Journal of Human–Computer Interaction*, 1–16. https://doi.org/10.1080/10447318.2022.2158263

Yao, F., & Wang, Y. (2020). Tracking urban geo-topics based on dynamic topic model. *Computers, Environment and Urban Systems*, *79*, 101419. https://doi.org/10.1016/j.compenvurbsys.2019.101419

Zhu, Z., Zhang, N., Ding, M., & Chen, L. (2023). Accountability mobilization, guanxi and social media-induced polarization: Understanding the bystander's prosocial punishment to misinformation spreader. *Information Systems Journal*, isj.12486. https://doi.org/10.1111/isj.12486

M-CATNAT: A Multimodal dataset to analyze French tweets during natural disasters

Badreddine Farah*
University of Orléans, INSA-CVL, LIFO, EA 4022, F45067 Orléans, France
badreddine.farah@univ-orleans.fr

Omar El Bachyr
University of Orléans, INSA-CVL, LIFO, EA 4022, F45067 Orléans, France
omar.elbachyr@usmba.ac.ma

Guillaume Cleuziou
University of Orléans, INSA-CVL, LIFO, EA 4022, F45067 Orléans, France
guillaume.cleuziou@univ-orleans.fr

Ana¨ıs Halftermeyerer
University of Orléans, INSA-CVL, LIFO, EA 4022, F45067 Orléans, France
anais.halftermeyer@univ-orleans.fr

Cécile Gracianne
BRGM, F45060 Orléans, France
c.gracianne@brgm.fr

Samuel Auclair
BRGM, F45060 Orléans, France
s.auclair@brgm.fr

Adel Hafiane
INSA-CVL, University of Orléans, PRISME, EA 4229, F18022 Bourges, France
adel.hafiane@insa-cvl.fr

Raphaël Canals
University of Orléans, INSA-CVL, PRISME, EA 4229, F45072 Orléans, France
raphael.canals@univ-orleans.fr

ABSTRACT

The proliferation of social media, especially platforms like X (formerly Twitter), has made available a large volume of real-time data valuable across diverse fields. During natural disasters, such data aids humanitarian efforts by providing crucial insights. However, processing this vast amount of data necessitates automated systems, often relying on annotated datasets for training. While supervised learning dominates this area, multilingual and multimodal annotated datasets are scarce. The present study addresses this gap by introducing M-CATNAT, a multimodal dataset of French tweets about natural disasters. Unlike previous datasets, M-CATNAT integrates annotations for texts, images, and their multimodal combination. Leveraging CrisisMMD guidelines, this work in progress aims to annotate 1,430 tweets, generating over 4,500 labels. The M-CATNAT dataset not only expands resources to non-English languages but also enhances multimodal analysis by furnishing three levels of annotation for each tweet (one per modality plus one for the whole tweet).

Keywords

Deep learning, French multimodal data, Crisis management, CrisisMMD Datatset

INTRODUCTION

The spread of social media use in this last decade has made available massive real time data. Particularly in platforms such as X (previously Twitter) where user-generated content has real value that can be used in a variety of applications including public health, economics and politics (Wu & Mebane Jr, 2022). When a natural disaster

*corresponding author

occurs, tweets can help humanitarian organizations to gather essential information in order to understand the scope of the disaster and prioritize actions to reduce suffering and rebuild communities (Alam et al., 2018). Nevertheless, the need to process the large amount of data available on social media requires the development of automated systems and methods that need annotated data to be trained.

The automation of social media information processing has recently become a hot research topic, with a majority of existing approaches relying heavily on supervised learning methods. Their primary objective is to categorize each post into a specific class based on the targeted task, such as informativeness (Alam et al., 2018) or relatedness (Kozlowski et al., 2020). Annotated datasets being the core element in supervised learning, considerable efforts have been made to manually construct such datasets. While most resources are text-based and predominantly in English, it's worth noting that resources in other languages also exist (Alharbi & Lee, 2019; Cobo et al., 2015; Cresci et al., 2015; Gru¨nder-Fahrer et al., 2018; Kozlowski et al., 2020) (see the following section for more details).

Despite numerous studies indicating that multimodality (e.g. text, images, video) can significantly enhance the performance of these methods (Abavisani et al., 2020; Ofli et al., 2020), there is currently only one extensively studied dataset – CrisisMMD (Alam et al., 2018), which focuses on multimodal English tweets related to crisis (natural disasters). Such a resource in English can be used to fined-tune a multilingual pre-training model (Conneau & Lample, 2019) and thus process French tweets. Although the preliminary results we have obtained with this process are not sufficiently satisfactory, they give reason to hope that the introduction of even a limited number of tweets in French during training would improve performance. Aligning the annotations of this new data (in French) with the annotations of the CrisisMMD dataset is, however, a prerequisite for success.

In this context, this work (in progress) aims to create the M-CATNAT dataset : a multimodal dataset of 1,430 French tweets about various natural disasters with annotations aligned with CrisisMMD tasks. However, the CrisisMMD dataset has the following limitation : it contains annotations for each of the two modalities of a tweet (text and image), but does not propose gold (i.e. manually annotated) annotations for the multimodal tweet as a whole (text+image). Since learning a multimodal model requires such a global label, tweets whose annotations for the two modalities do not match are generally filtered out (Ofli et al., 2020). Unlike CrisisMMD, the M-CATNAT dataset we are preparing includes not only gold annotations for the image modality and the text modality but also the gold label of the tweet as a whole, i.e the multimodal instance (text+image). Interestingly, the full annotation model proposed makes it possible to offer a greater representativeness of tweets in terms of the diversity of relationships between modalities (e.g. redundancy, complementary or contradiction in the information).

In this article, we present the framework defined for the multimodal annotation process of the 1,430 French tweets that will make up the M-CATNAT dataset[1].

So far, seven annotators from the research team have fully annotated 837 tweets, out of the 1,430 planned. This ongoing process has resulted in the creation of more than 3,000 manual annotations, as each instance entails three labels (one for the image, one for the text, and one for the multimodal instance with some tweets having multiple images). The CrisisMMD humanitarian classes definition guided our annotation process, ensuring consistency and reliability in the labeled data. In addition, the annotation process was regularly evaluated to check both its consistency (inter-annotator agreement) and its alignment with the CrisisMMD dataset.

RELATED WORKS

The surge in the use of social media has led to an increased interest in processing their content with numerous studies focusing, in particular, on crisis-related tweets. As reported in Table 1, a lot of emphasis has been placed on text-based analysis, given that the majority of annotated datasets primarily feature labels on text alone. These datasets have introduced various tasks, such as assessing the *relatedness* of tweets to a specific disaster (Kozlowski et al., 2020), determining the *informativeness* of tweets (Alam et al., 2021; Olteanu et al., 2015), classifying *crisis types*, and detecting *eyewitnesses* (Zahra et al., 2020). In the domain of image processing for tweets, tasks include estimating the *damage severity* (Alam et al., 2018; Nguyen et al., 2017) and *damage type* (Mouzannar et al., 2018), evaluating *informativeness* (Alam et al., 2018). Most resources are in English, although resources also exist in other languages like Spanish (Cobo et al., 2015), Italian (Cresci et al., 2015), German (Gru¨nder-Fahrer et al., 2018) and Arabic (Alharbi & Lee, 2019). For the French language, (Kozlowski et al., 2020) proposed a text only disaster tweet dataset with three classification levels including *relatedness*, *urgency* and *intent to act*. One should note that the various tasks addressed lead the community to produce datasets more or less compatible with the central resource that is CrisisMMD. Each dataset whose task mentions *humanitarian* is either an implementation of the CriseMMD guidelines or a mapping to a simplified set of classes, keeping it compatible with CrisisMMD[2].

1The dataset is available here : https://github.com/badreddineFarah/M-CATNAT

2For a substantial survey of existing datasets for English, see (Feng et al., 2022).

Table 1. Crisis related published resources.

	Modal. labels	Tasks	Platform	Size	Language
(Olteanu et al., 2015)	Text	informativeness, humanitarian, source	Twitter	28,000	English
(Alam et al., 2021)	Text	informativeness, humanitarian	Twitter	166,098 (inf.), 141,533 (hum.)	English (94%)
(Kozlowski et al., 2020)	Text	relatedness, urgency, intent to act	Twitter	12,826	French
(Cobo et al., 2015)	Text	relevancy	Twitter	2,187	Spanish
(Cresci et al., 2015)	Text	relatedness, damage	Twitter	5,642	Italian
(Alharbi & Lee, 2019)	Text	informativeness, humanitarian	Twitter	4,037	Arabic
(Zahra et al., 2020)	Text	eyewitness	Twitter	14,000	English
(Alam et al., 2023)	Image	damage severity, humanitarian, disaster type	Twitter, Google, ...	71,198	-
(Nguyen et al., 2017)	Image	damage severity	Twitter, Google	25,758	-
(Alam et al., 2018)	Image, text	informativeness, humanitarian, damage severity	Twitter	16,097	English
(Mouzannar et al., 2018)	Multimodal	damage type	Instagram, Twitter, Google	5,879	English
M-CATNAT (Ours)	**Image, text, multimodal**	**Informativeness, humanitarian**	**Twitter**	**1,430** (in progress)	**French**

Table 2. Disaster events considered.

Crisis type	Affected areas	Search start date	Search end date	# collected tweets	# sampled tweets (images)
Floods	Alpes-Maritimes et Var	10/2/2015 12:00:00 AM	10/4/2015 11:59:00 PM	14846	108 (124)
	Carcassonnais	10/14/2018 12:00:00 AM	10/15/2018 11:59:00 PM	14341	171 (205)
	Secteurs de Be´ziers / Narbonnais	10/22/2019 12:00:00 AM	10/24/2019 11:59:00 PM	7550	104 (126)
	Pays Basque / Be´arn/ Pyre´ne´es	12/12/2019 12:00:00 AM	12/15/2019 11:59:00 PM	6838	132 (158)
	Cote d'Azur 83/06	11/22/2019 12:00:00 AM	11/24/2019 11:59:00 PM	9983	164 (206)
	Cote d'Azur 83/06	12/1/2019 12:00:00 AM	12/1/2019 11:59:00 PM	6353	89 (103)
	Valleraugue, Saumane	9/19/2020 12:00:00 AM	9/20/2020 11:59:00 PM	5721	72 (83)
	Alex storm (south of France)	10/1/2020 12:00:00 AM	10/3/2020 11:59:00 PM	7804	110 (129)
		Date	Time		
Earthquakes	Barcelonnette	07/04/2014	07:26:59 PM	11085	66 (72)
	La Rochelle	28/04/2016	06:46:53 AM	3802	87 (91)
	Thouars	21/06/2019	06:50:57 AM	2668	56 (58)
	Le Teil	11/11/2019	10:52:45 AM	6448	192 (224)
	Strasbourg	12/11/2019	01:38:00 PM	3123	79 (86)
Total				100562	**1430 (1665)**

Despite recent studies highlighting the effectiveness of multimodal datasets in improving the performance of machine learning models through multimodal training (Abavisani et al., 2020; Liang et al., 2022; Long & McCreadie, 2022; Ofli et al., 2020), limited attention has been given to the creation of new multimodal crisis-related datasets and CrisisMMD remains the most widely used dataset despite its inherent limitations. A first notable limitation is that CrisisMMD independently annotates images and text, necessitating practitioners to filter instances with discordant labels (different labels for image and text) and retain instances with consistent labels for training. This approach does not take into account the various relationships between modalities in social media data (Vempala & Preo¸tiuc-Pietro, 2019). In response to this gap, (Sosea et al., 2021) introduced Disrel, a dataset for classifying the relationship between image and text. Their work demonstrated that incorporating such tasks can enhance the performance of models trained on CrisisMMD.

This study aims to enhance CrisisMMD in two dimensions. Firstly, we augment the corpus by incorporating French tweets, enabling multilingual and multimodal training. Secondly, we provide three annotations for each instance, hypothesizing that this enriched dataset will enable models to better capture the relationships between images and texts, thereby improving overall system performance. Importantly, we adhered to the CrisisMMD guidelines to ensure alignment between our annotations and the original annotation guide.

DATA COLLECTION

To initiate the annotation process, it is essential to gather data. For this purpose, we used datasets collected as part of the Re´SoCIO project by the French Geological Survey (BRGM) via its SURICATE-Nat platform (Auclair et al., 2019). The data collection process relies on the interrogation of Twitter's "academic" API on the basis of a carefully chosen keyword search, exploring French-speaking lexical fields related to floods on the one hand, and earthquakes on the other. Below are the specific queries used for each type of disaster, with the aim of retrieving a maximum number of posts describing the effects of the phenomena, as well as a minimum number of off-topic posts.

Flood key-words The keywords used are the following (taking spelling errors into account) : "*inondation*" (flood), "*inonde´*" (flooded / inundated), "*sous l'eau*" (under water), "*rivie`re en crue*" (swollen river), "*crue*" (flood upward trend / freshet), "*de´crue*" (flood downward trend), "*onde de crue*" (flood wave), "*sort de son lit*" (rise above the banks), "*torrentiel*" (torrential), "*emporte´ par les eaux*" (washed away). The query excludes retweets and tweets with flood lexical fields related to politics, sex or migration, while focusing only on French language tweets by specifying "lang:fr".

Earthquake key-words "*Se´isme*" or "*tremblement de terre*" (earthquake), "*magnitude*" (magnitude), and "*terre tremble*" (ground shaking) are the keywords used to retrieve the tweets (taking spelling errors into account as well). The query excludes retweets and tweets with earthquake lexical fields related to sex, migration or politics, while focusing only on French language tweets by specifying "lang:fr".

DATA SAMPLING

As illustrated in Table 2, we sampled tweets from various events of natural disaster, taking one-third from earthquake events and two-thirds from flood events, in order to respect the representativeness of each disaster while keeping the number of tweets sufficient for each disaster. The earthquakes and flash floods selected correspond to significant

events for mainland France, from a phenomenological and/or a crisis management point of view. To maintain data quality and relevance, we applied additional criteria, exclusively selecting tweets with text lengths between 5 to 40 words to ensure sufficient textual content for comprehensive analysis. Additionally, we selectively included tweets with at least one associated image, as the presence of visual content is pivotal for preserving the multimodal nature of the dataset.

The goal is to annotate a total of 1,430 multimodal tweets, with each tweet divided into three instances: text alone, image alone, and the combination of text and image, resulting in approximately 4,775 instances for annotation (because some tweets contain several images). For tweets with multiple images, we annotate the text, each image, and multiple text-image combination instances are created by merging the tweet's text with each associated image.

It is worth noting that since the beginning of the annotation process, around 5% of the data have been deleted from X, leading us to revise the dataset size from 1,500 to 1,430 tweets. In addition, it is still possible that between now and the end of the annotation process, new tweets will be deleted, reducing the total number of tweets available for annotation.

TASKS DESCRIPTIONS

As said before, we use the *humanitarian* and *informative* tasks (from CrisisMMD) to annotate the dataset[3]. However, we also adapt the task to the practises in the community, as its was already done within several works (Abavisani et al., 2020; Liang et al., 2022; Long & McCreadie, 2022; Ofli et al., 2020) which have merged classes relative to damages ("infrastructure_and_utility_damage" and "vehicle_damage") into "infrastructure_and_utility_damage" and classes relative to human casualties ("affected_individuals","injured_or_dead_people","missing_or_found_people") into "affected_individuals", resulting in a five classes task described bellow:

Affected individuals If the tweet/image reports or shows individuals affected by the disaster event, such as people sitting outdoors, individuals standing in lines for assistance, people in need of shelter facilities, missing or found individuals, or deceased individuals.

Infrastructure and utility damage If the tweet/image reports or shows a damaged structure or one whose use is affected by an earthquake, fire, heavy rains, floods, strong winds, gusts, etc., such as damaged houses, roads, buildings; flooded houses, streets, highways; blocked roads, bridges, paths; collapsed bridges, power lines, cars, boats, communication poles, etc.

Rescue volunteering or donation effort If the tweet/image reports or shows any type of rescue, volunteering, or donation effort, such as transporting people to safe locations, evacuating people from the hazardous area, individuals receiving medical or food aid, people in shelter facilities, monetary donations, blood donations, or services, etc.

Other relevant information If the tweet/image does not fit into any of the three categories above but still contains important information useful for humanitarian aid.

Not related or not relevant information If the tweet/image is not useful for humanitarian aid.

As described in the CrisisMMD paper, only informative tweets have been annotated into one of the *humanitarian* classes. Thus, considering other tweets as not informative (not related or not relevant) the annotated dataset can also be used to train models on the task *informativeness*.

ANNOTATION METHODOLOGY

In the annotation methodology, we implemented a four-phase process to ensure high-quality annotation of the dataset. At the end of each phase, an assessment of the annotations was made according to different indicators, followed by a consultation meeting focusing on examples with substantial disagreement occurred. A clarification, accompanied by illustrative examples was then incorporated into the annotation guide. Also, to mitigate the bias in labelling the modalities from each (multimodal) tweet, the order of presentation of the instances to be annotated needs to be carefully considered. An annotator who is first presented with a complete multimodal tweet (text +

[3]The annotation guide is delivered with the resource (https://anonymous.4open.science/r/M-CATNAT-5692/README.md).

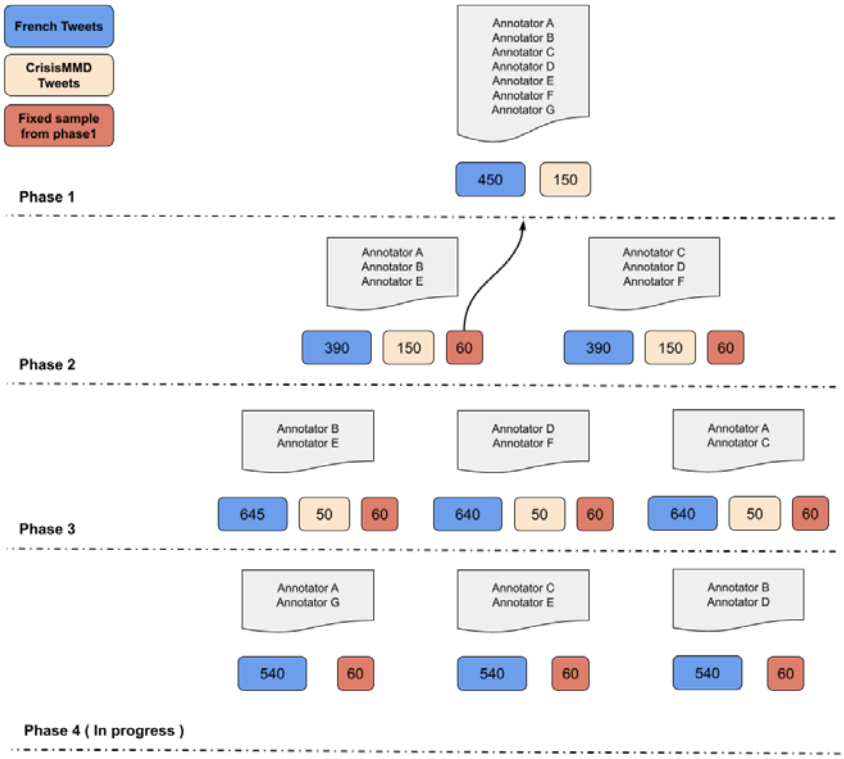


Figure 1. Overview of the annotation protocol. This diagram illustrates our multi-phase annotation methodology. In each phase, we annotate a specific number of new instances (image, text and multimodal instances), highlighted in blue. To maintain alignment with the CrisisMMD, a small portion of this dataset is also annotated. Lastly, to ensure the self-consistency of each annotator, some instances are selected for re-annotation.

image) will undoubtedly be influenced in his annotation (a posteriori) of each of the modalities. We have therefore ensured that the text and image modalities are presented first and in a separate slot (groups of annotators) or phases from the multimodal tweet (text+image).

This sequential presentation aimed to ensure independent labeling decisions for both unimodal and multimodal instances, minimizing potential bias in annotations.

As presented in Figure 1, Phase 1 involved seven annotators collectively annotating one shared dataset of 600 instances constituted of 150 instances from CrisisMMD and 450 instances from new French tweets ($\frac{1}{3}$images, $\frac{1}{3}$ texts and $\frac{1}{3}$ multimodal instances). In this phase, when an instance is annotated identically by at least 5 annotators (out of 7), this annotation is definitively used as a label. The other instances (for which the annotation is not consensual) are discussed and collectively labelled at the end of the phase during the consultation meeting. This process allowed us to clarify and adjust the annotation guide. This first phase is an agreement (or training) phase for the annotators. Subsequent phases gradually "industrialise" the annotation process, with annotators organised into parallel pools.

Phase 2 comprised two pools, each pool was made up of 3 annotators responsible for annotating 600 instances. Among the 600 instances, we maintained a set of 150 CrisisMMD instances and introduced 60 instances from phase 1 in order to compute the consistency of annotations through the phases for each annotator involved in the process. At the end of this phase, the instances with two identical annotations are definitively labelled, while we proceed to a correction phase during which each disagreement is resolved by an annotator from the other pool.

In phase 3, we decomposed our team into three annotation pairs, for each annotation pool we reduced the number of CrisisMMD instances to 50 while keeping the 60 instances from phase 1. The disagreement is resolved by a third annotator from another pool.

Currently ongoing, Phase 4 involves three new pairs of annotators, each annotating 540 instances, along with the 60 instances from Phase 1. The ultimate goal is to annotate the full set of 1,430 tweets, amounting to approximately 4,775 manual annotations at the conclusion of the process.

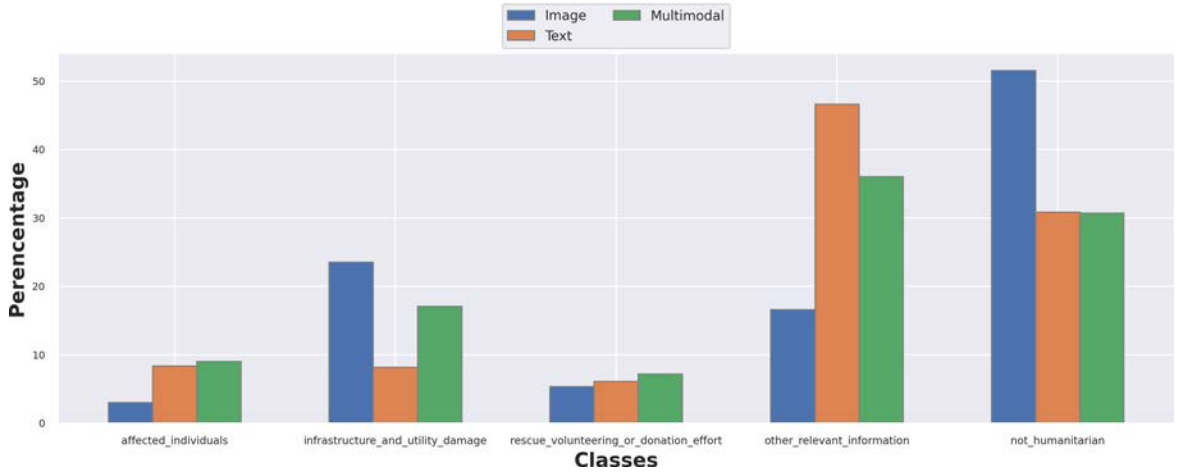


Figure 2. Comparison of label distribution for each modality, the distributions are calculated on a 1,032 fully annotated instances.

RESOURCE ANALYSIS

In this section, we provide a comprehensive analysis of the M-CATNAT resource at the time of writing (837 unique tweets fully labelled corresponding to 1,032 multimodal pairs since some tweets have multiple images). Firstly, we outline the distribution of labels for each modality and for each type of disaster (earthquakes and floods). Secondly, we present the results of annotation scores, including the Fleiss Kappa score (Fleiss et al., 2013) and the CrisisMMD alignment score. Finally, we show the distribution of class concordance/discordance according to modalities.

LABEL DISTRIBUTION

Figure 2 illustrates the label distribution across different classes. Notably, the class "Affected Individuals" exhibits the lowest representation. The "Not Humanitarian" class, equivalent to "Not informative" in the *informativeness* task

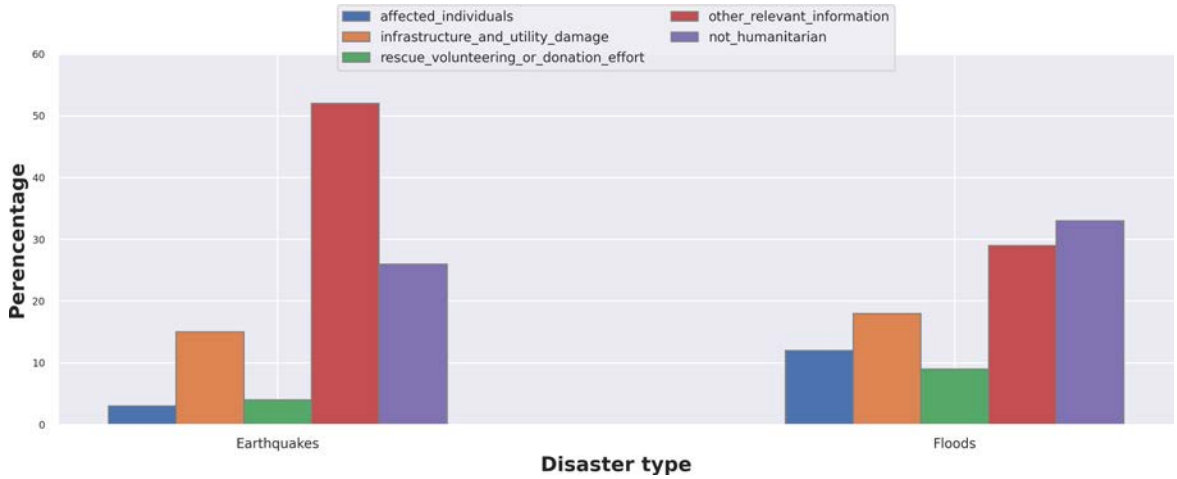


Figure 3. Comparaison of label distribution for earthquakes and floods, the result are on 307 earthquake related tweets and 725 floods related tweets.

Table 3. Fleiss kappa scores on each phase.

Phase	1	2	3
Mean (over pools)	0.64	0.70	0.61
Common examples	0.64	0.68	0.66

Table 4. CrisisMMD alignment scores on each phase.

Phase	1	2	3
CrisisMMD alignment	79%	74%	77%

of CrisisMMD, constitutes 30% in text and multimodal instances but rises to approximately 50% in image labels. Conversely, "Other Relevant Information" represents 46% in texts and 36% in multimodal instances, contrasting with its lower representation at 16% in images. This difference is attributed to the enriching role of textual information in multimodal instances, aiding in disaster characterization regarding intensity and location. The distribution of multimodal labels closely aligns with the text distribution, highlighting the importance of textual information in crisis datasets in general and in the M-CATNAT dataset in particular. The emphasis on textual information stems from its reflective of tweet intentions and provision of more precise details. Classes such as "Rescue Volunteering and Donation Efforts" are more prominent in CrisisMMD (14.5%), indicating that the disasters in our data are comparatively less severe.

Figure 3 reports the distribution of multimodal annotations according to the type of disaster in our dataset. "Other Relevant Information" exhibits higher representation in earthquake-related tweets, due to the lower magnitude of these earthquakes where tweets often provide basic information about the incident, such as location and magnitude. Conversely, "Affected Individuals" and "Rescue volunteering or donation effort" are more represented in flood-related tweets, reflecting the extensive damage caused by such events.

ANNOTATOR SCORES

In accordance with the annotation methodology outlined in the "Annotation Methodology" section, our process involves four phases, each subdivided into annotator pools. In consequence we propose to measure inter-annotator agreement on each phase using the Fleiss Kappa (Fleiss et al., 2013)[4]. Table 3 reports the results obtained: first row shows the mean Fleiss Kappa score overs pools for each phase, while in the second row Fleiss Kappa scores were calculated specifically for the 60 common examples (highlighted as red samples in Figure 1). The variations observed are mainly due to the changes made to the annotation guide over the phases; it is also significant to note that the inter-annotator agreements remained above 0.60 in each of the three phases. To be complete, we noticed the agreement between annotators varied across modalities. A higher agreement is observed for images (up to 0.85) compared to texts which exhibited a poorer annotator agreement, particularly in Phase 1 (0.56).

4In addition, at every phase we made annotations for a sample of 60 fixed instances (in red in Figure 1) in order to monitor the intra-annotator agreement along the whole process. The percentage scores obtained vary between 67% and 78% for the transition from phase 1 to phase 2 and between 82% and 89% for the transition from phase 2 to phase 3.

CRISISMMD ALIGNMENT

Moreover, we assessed alignment with the CrisisMMD dataset (see Table 4). In Phase 1, alignment was computed on 150 examples after disagreement resolution, yielding a score of 79% (with a notable increase to 90% for multimodal instances). In Phase 2, the alignment score was computed on 150 examples for each pool and the mean score obtained is 74%, while it reached 77% in phase 3 on 50 examples for each pool (mean score of the three pools is reported). These results are encouraging given our will to use the proposed resource in addition to CrisisMMD in a multilingual configuration.

MODALITY LABEL ANALYSIS

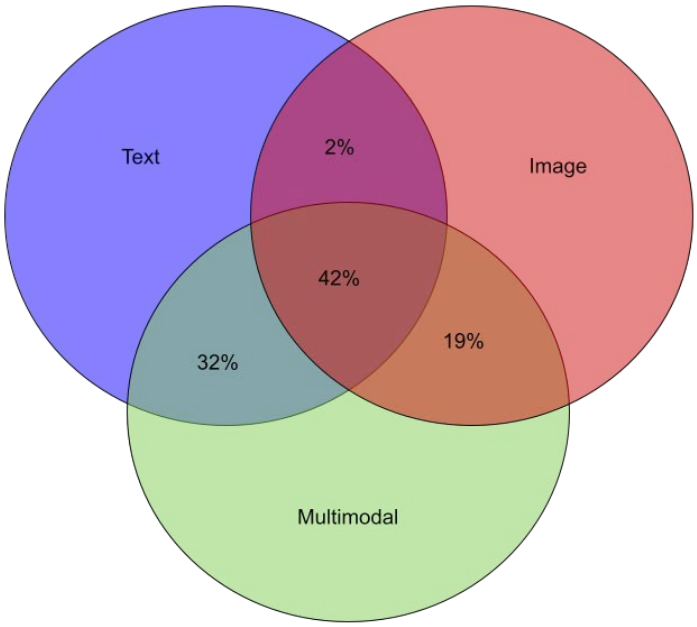


Figure 4. Label Combinations for Image, Text, and Multimodal on 1,032 fully annotated instances.

Figure 4 shows an analysis of the labelling of the 1,032 5 examples in the current M-CATNAT dataset. It reveals the relation in the distribution of labels across the three labels assigned to one Tweet : the *image* label, the *text* label and the *multimodal* label. A significant portion, accounting for 42% of the examples, demonstrates consistency across all three labels ; this suggests a strong correlation among the different types of data and their respective labels. In 32% of cases, although the text label differs from the image label, it aligns with the multimodal label ; this implies that the textual content plays a crucial role in shaping the combined interpretation. Conversely, 19% of the examples have an image label distinct from the text label but aligning with the multimodal one, highlighting the influence of visual elements in these instances. Following these results, one may consider the need to train models to spot instances where the information looked for may be found in a specific modality. Lastly in a small proportion, just 2%, the image and text labels match but differ from the multimodal label, and lastly, 5% of cases shows complete disparity across all three labels. These findings underscore the variety of relations between different modalities and emphasize the importance of considering multimodal annotation rather than an independent annotation of the text and the image.

It's worth noting that crisisMMD use cases typically focus only on instances with identical labels across modalities. However, our study overlooks more than 50% of the data, disregarding the diverse relationships between images and text present in the multimodal tweets dataset.

CONCLUSION

Leveraging social media information during natural disasters can significantly help humanitarian aid organizations. Deep learning models play a crucial role in filtering and classifying this data efficiently. However, the effectiveness of such models relies heavily on labeled data for training. Research has highlighted the potential of multimodal

5We can get 1,032 image-text instances from the 837 annotated tweets as a tweet can have multiple images.

data in enhancing performance, yet available datasets are predominantly unimodal and in English. Although one multimodal dataset exists, its separate labeling for images and text poses limitations. This work addresses this gap by introducing M-CATNAT, a French multimodal dataset aligned with CrisisMMD, featuring three labels for each instance. This alignment facilitates multilingual training, thereby broadening the applicability of deep learning techniques. We provide annotator scores and other analyses to support deep learning practitioners in utilizing the dataset effectively. In future work, our aim is to further expand the annotated data and train models on this dataset.

ACKNOWLEDGEMENT

This work was supported by the French national research agency (ANR) under the program IA.iO (ANR-20-THIA-0017-01), as well as within the Re´SoCIO project (ANR-20-CE39-001). Opinions expressed in this paper solely reflect the authors' view; the ANR is not responsible for any use that may be made of information it contains. The computation was performed using Leto resources from CASciModOT federation.

REFERENCES

Abavisani, M., Wu, L., Hu, S., Tetreault, J., & Jaimes, A. (2020). Multimodal categorization of crisis events in social media. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14679–14689.

Alam, F., Alam, T., Hasan, M. A., Hasnat, A., Imran, M., & Ofli, F. (2023). Medic: A multi-task learning dataset for disaster image classification. *Neural Computing and Applications*, *35*(3), 2609–2632.

Alam, F., Ofli, F., & Imran, M. (2018). Crisismmd: Multimodal twitter datasets from natural disasters. *Twelfth international AAAI conference on web and social media*.

Alam, F., Sajjad, H., Imran, M., & Ofli, F. (2021). Crisisbench: Benchmarking crisis-related social media datasets for humanitarian information processing. *Proceedings of the International AAAI Conference on Web and Social Media*, *15*, 923–932.

Alharbi, A., & Lee, M. (2019). Crisis detection from arabic tweets. *Proceedings of the 3rd workshop on Arabic corpus linguistics*, 72–79.

Auclair, S., Boulahya, F., Birregah, B., Quique, R., Ouaret, R., & Soulier, E. (2019). Suricate-nat: Innovative citizen centered platform for twitter based natural disaster monitoring. *2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*, 1–8.

Cobo, A., Parra, D., & Navo´n, J. (2015). Identifying relevant messages in a twitter-based citizen channel for natural disaster situations. *Proceedings of the 24th international conference on world wide web*, 1189–1194.

Conneau, A., & Lample, G. (2019). Cross-lingual language model pretraining. *Advances in neural information processing systems*, *32*.

Cresci, S., Tesconi, M., Cimino, A., & Dell'Orletta, F. (2015). A linguistically-driven approach to cross-event damage assessment of natural disasters from social media messages. *Proceedings of the 24th International Conference on World Wide Web*, 1195–1200.

Feng, Y., Huang, X., & Sester, M. (2022). Extraction and analysis of natural disaster-related vgi from social media: Review, opportunities and challenges. *International Journal of Geographical Information Science*, *36*(7), 1275–1316.

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical methods for rates and proportions*. john wiley & sons.

Gru¨ nder-Fahrer, S., Schlaf, A., Wiedemann, G., & Heyer, G. (2018). Topics and topical phases in german social media communication during a disaster. *Natural language engineering*, *24*(2), 221–264.

Kozlowski, D., Lannelongue, E., Saudemont, F., Benamara, F., Mari, A., Moriceau, V., & Boumadane, A. (2020). A three-level classification of french tweets in ecological crises. *Information Processing & Management*, *57*(5), 102284.

Liang, T., Lin, G., Wan, M., Li, T., Ma, G., & Lv, F. (2022). Expanding large pre-trained unimodal models with multimodal information injection for image-text multimodal classification. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15492–15501.

Long, Z., & McCreadie, R. (2022). Is multi-modal data key for crisis content categorization on social media? *ISCRAM 2022 Conference Proceedings 226 19th International Conference on Information Systems for Crisis Response and Management*, 1068–1080.

Mouzannar, H., Rizk, Y., & Awad, M. (2018). Damage identification in social media posts using multimodal deep learning. *ISCRAM*.

Nguyen, D. T., Ofli, F., Imran, M., & Mitra, P. (2017). Damage assessment from social media imagery data during disasters. *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, 569–576.

Ofli, F., Alam, F., & Imran, M. (2020). Analysis of social media data using multimodal deep learning for disaster response. *arXiv preprint arXiv:2004.11838*.

Olteanu, A., Vieweg, S., & Castillo, C. (2015). What to expect when the unexpected happens: Social media communications across crises. *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, 994–1009.

Sosea, T., Sirbu, I., Caragea, C., Caragea, D., & Rebedea, T. (2021). Using the image-text relationship to improve multimodal disaster tweet classification. *The 18th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2021)*.

Vempala, A., & Preo¸tiuc-Pietro, D. (2019). Categorizing and inferring the relationship between the text and image of twitter posts. *Proceedings of the 57th annual meeting of the Association for Computational Linguistics*, 2830–2840.

Wu, P. Y., & Mebane Jr, W. R. (2022). Marmot: A deep learning framework for constructing multimodal representations for vision-and-language tasks. *Computational Communication Research*, *4*(1).

Zahra, K., Imran, M., & Ostermann, F. O. (2020). Automatic identification of eyewitness messages on twitter during disasters. *Information processing & management*, *57*(1), 102107.



Impact-driven Emergency and Crisis Management

Track Chairs

Melanie Reuter-Oppermann, University of Twente

Laura Boyle, Queen's University Belfast

Christian Elsenbast, Fraunhofer IESE

State of Knowledge Management among German firefighters

Andreas Schultz*
Paderborn University
Data management in mechanical engineering (DMB)
andreas.schultz@uni-paderborn.de

Fabian Dotzki
Paderborn University
Data management in mechanical engineering (DMB)
dotzki@mail.uni-paderborn.de

Iryna Mozgova
Paderborn University
Chair of Data management in mechanical engineering (DMB)
iryna.mozgova@uni-paderborn.de

ABSTRACT

This paper introduces the results of a survey that was conducted among German firefighters in 2023 to investigate the current situation of knowledge management. We want to examine which systems and techniques are in use, what the end users' expectations towards knowledge management are, and which requirements a system should fulfill. We focus on all stages of the crisis management circle, from prevention and preparation to the response to an incident as well as recovery.

The goal of this survey is to evaluate firefighters' needs in order to be able to define requirements for a knowledge management system that will be accepted by the users. The exchange of information items between entities (emergency units, authorities, and enterprises) and other IT systems is one important aspect. We identify existing information and knowledge management systems, and means to improve user satisfaction by leveraging users to maintain a system with relevant and up-to-date information and knowledge.

Keywords

Knowledge management, firefighter, emergency management, survey, Germany

INTRODUCTION

During the past decades, crisis management has experienced many changes regarding exterior circumstances and interior preconditions. New and changing legal and administrative regulations hamper the performance of emergency management organizations, digitization puts new requirements on the practitioners' skills (cf. (Sauerland et al., 2020)) and the technical and procedural understanding has to grow constantly. It is challenging to find new (voluntary) personnel as their expectations shift towards work- and family-friendly conditions and the level of skills and knowledge required continues to increase. Since the amount of necessary time is constantly increasing, measurements have to be taken into account to cope with it. There are few publications related to the combination of knowledge management and crisis management, as stated by Anand et al. in their survey study related to knowledge management in contexts of crisis (Anand et al., 2022). They identified 59 publications focusing on knowledge management and disaster management, but factors supporting the development of knowledge management for crisis preparedness is scant. Oktari et al. showcase the positive effects of knowledge management that improves both the effectiveness and performance of operations as well as innovation among users (Oktari et al., 2020). Only few literature can be found that deals with the actual implementation of a knowledge management system for emergency

*corresponding author

services. One older article from Pickles describes the development of a knowledge management infrastructure in Australia but hardly more information about the system can be found (Pickles, 2004). Another description comes from Timm et al. who describe the comparison of different open-source systems to be used as a knowledge and collaboration platform (Timm et al., 2013). Some German fire brigades make use of the Einsatzleiterwiki, which is a software solution that stores incident-relevant information in a wiki. A public instance exists that is accessible on the internet[1]. The usage as a closed system by the Berlin fire department is described in (Ziehr et al., 2017). This shortage in knowledge management research led the authors to perform a survey among German firefighters to find out what their attitudes towards knowledge and knowledge management are.

This survey is one of the first steps in a holistic process of examining knowledge management among firefighters, surveying the current status and needs of emergency personnel regarding knowledge management, and finally developing a system that incorporates the findings. It aims to bridge the gap between research and practice by laying the foundation to develop a system according to the needs of end users. The focus is on the German firefighters, where each federal state (Bundesland) has to govern the emergency management and issue laws and regulations (except civil protection for which the Federal State (Bund) is responsible for) on its own. The vast majority are voluntary firefighters, and in general, the bigger the city, the more professional firefighters are on duty, mostly 24/7. See (Deutscher Feuerwehrverband, 2021) and Figure 6 for details. The results of the survey should be independent of the specific organisation of the participants. Of course, some organisations either have more personal capacities like professional personal who are employed or have motivated voluntary members who are already involved in maintaining a knowledge management systems (KMS). In general, the prerequisites should be similar between fire services, relief organisations or other emergency organisations and a comparability of the survey results is given. In contrast to KMS in a corporate environment, KMS for emergency services can profit from an absent competitive situation which would impede the exchange of information between organisations.

In this evaluation, a methodical approach was followed, guided by the methods recommended in *Umfragen erstellen und auswerten* (Creating and Evaluating Surveys) (Ortmanns & Sonntag, 2023). The central research question is: To what extent is knowledge management bruited in German fire departments, particularly outside of emergency situations, and what are the end users' expectations towards a KMS. The extent of the dissemination and intensity of the use of KMS were examined. The hypothesis was formulated that the spread and use of such systems in German fire departments is currently limited and confined to a few projects. Emergency personnel, although not stated explicitly in literature, wishes for a system to manage their knowledge.

The survey plan was developed in line with the methods suggested by (Ortmanns & Sonntag, 2023). The investigation employed a quantitative survey method, supplemented by the inclusion of some open questions to allow for a broader range of responses. The survey was tested before dissemination internally several times and by few senior firefighters known to the authors. Particular care was taken to reach as many firefighters as possible. To this end, fire department associations in Germany were contacted, and the survey was distributed via social media. The data collection was conducted online with the EU Survey system[2] to make use of a flexible and proven system to ensure maximum reach and user friendliness.

Terms and concepts related to knowledge management often do not have a universally valid definition. Therefore, we provide a disambiguation and distinguish between the terms data, information, and knowledge in the following way, according to (North, 2016; Mozgova et al., 2020) and depicted in Figure 1:

- Symbols are the basic set of available characters (letters, numbers) that, when combined in a syntactical way, form data.
- Data are uninterpreted sets of symbols that do not have a specific meaning until a sense is added.
- Information is data with a context of meaning and an added relationship.
- Knowledge is the result of connected information. Information forms the basis and needs to be interpreted by a human being.

The remaining stairs in 1 refer to a more business-oriented view on knowledge management, which is only partly applicable for emergency management (Duscher, 2010). At least the last step, competitiveness, does not make sense for emergency personnel since there is no need to compete rather than to collaborate.

This paper is structured as follows: At first, we will give an overview of the survey and present our methodology as well as demographic data about our respondents. Then we will examine the knowledge management-related

1https://einsatzleiterwiki.de/
2https://ec.europa.eu/eusurvey/

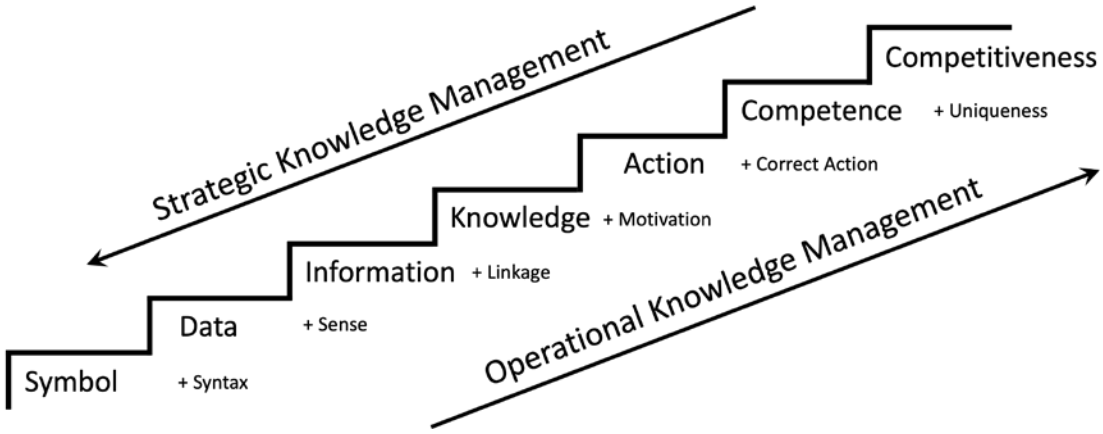


Figure 1. Knowledge staircase according to (North, 2016)

aspects of our survey, where we elaborated on the current status of German fire brigades as well as respondents' requirements and wishes. Next, we present the findings of our survey and deduce aspects that have to be taken into account when developing a KMS for emergency personnel. Finally, a conclusion summarizes our presentation and shows further areas that need to be examined.

SURVEY

The survey was designed to get an overview of current aspects related to knowledge and knowledge management among firefighters, from the beginner to senior experts with decades of experience. The procedure for the whole survey is shown in Figure 2. During creation, the questionnaire was divided into the blocks *statistical data* (8 questions), *information management* (19 questions plus free text fields to clarify answers), *information exchange* (5 questions plus free text fields to clarify answers), and *data maintenance* (21 questions plus free text fields to clarify answers). The survey period was from August 2023 until October 2023. We distributed the survey among both, federal and state fire organizations and via personal contacts with the request to forward it to them, and we posted calls to encourage participation online via Facebook posts and groups. We received a total of 577 responses from firefighters across Germany. Only 5 submissions were empty and excluded beforehand. Since some questions were only shown based on certain previous answers and other questions have multiple answers possible, the number of responses per question might vary. The evaluation is carried out in this paper.

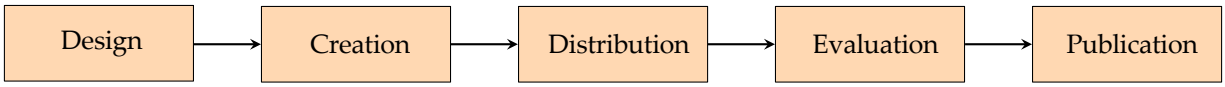


Figure 2. Survey process

We contacted each state fire association (Landesfeuerwehrverband) in Germany but only two replied positively; one declined our request, and the rest did not answer at all. The distribution via Facebook by the DFV[3] is likely location-independent, other invitations to participate via Facebook groups might be filtered more restrictively by Facebook's algorithms. The participants and results of our survey roughly align both with the population's distribution in Germany, as indicated in Figure 3, as well as the distributions of education and expertise among the firefighters. But we do not consider our study fully representative for Germany since the emergency domain differs in each state and we do not have sufficient answers from every state.

Overview

Beginning with demographic data for the subsequent analysis, we asked the participants questions to get an idea of their type and location of engagement, experience in the emergency domain, their IT and leadership skills, and their age and duration of membership. In the following, we considered the current state of information and knowledge

[3] https://www.facebook.com/112willkommen/posts/pfbid02kjfzQVu9T5UNjajxDo9qQ7zQiaheFCE8nAYgqzvMfgBJimWRhp7GAxVc5qmFX5Qjl retrieved August 2, 2023

management in their units. We were interested in exploring how information and knowledge are currently being exchanged among firefighters and the tools they utilize for this purpose. We aligned our survey with information from the DIN ISO 30401 norm, which deals with requirements for KMS in business contexts (Deutsches Institut fu¨r Normung, 2022). Another important step in managing knowledge is the exchange of already-available information via Information Management Systems (IMS), respectively KMS. Therefore, we examined the distribution of means to share data between entities (emergency units, authorities, and enterprises) or IT systems. In the last section, we examined the current status and willingness to maintain KMS. We wanted to see whether our participants are already involved in creating knowledge, instructions, or sharing information and how they do it. We asked them to rate several statements regarding the upkeep of knowledge in a KMS, focusing on the amount of time they devote or would devote to maintenance tasks and their preferred tools for these activities.

Detailed Insights

We received a total of 577 valid answers that were taken into account for our validation. Only one answer is not based within Germany; five did not provide their origin. The federal states of North Rhine-Westphalia and Lower Saxony are particularly well represented. City states like Berlin, Hamburg, or Bremen have the lowest numbers of responses. The distribution of responses by state, as shown in Figure 3, does not perfectly match the distribution of the German population across federal states. However, our results are generally in line with the population distribution. A comparison between the actual numbers of firefighters per state with the submitted surveys would have been preferable, but it is difficult to obtain comparable recent numbers for all federal states.

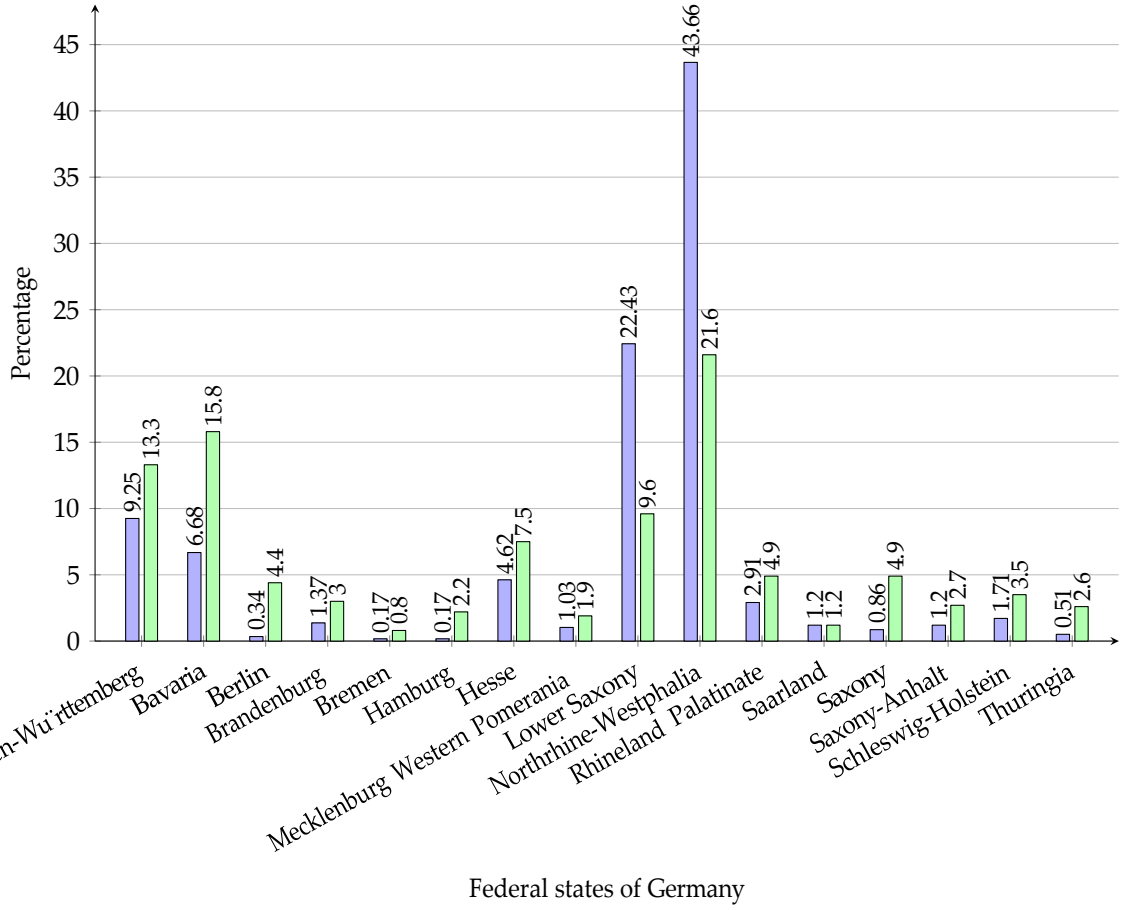


Figure 3. Distribution of survey answers (in blue) compared to inhabitants (in green), data from (Bundeszentrale fu¨r politische Bildung, 2023)

The age structure of our participants shows a relatively even distribution across all ages, ranging from 18 to the age at which emergency personnel typically retire. According to the German fire brigade association (Deutscher Feuerwehrverband), the age limit for active duty in emergency services varies from state to state. For instance, in some states, young people can participate with limitations beginning at 16 years old, while in other states, the minimum age is 17 or 18 years old. The upper age limit also shows variation, ranging between 60 and 67 years across different regions (Deutscher Feuerwehrverband, 2017).

The duration of membership of our respondents appears to be evenly distributed across the 10-year categories. Significantly, a number of participants with decades of experience contributed to our survey, providing valuable insights into their extensive knowledge, as depicted in Figure 4.

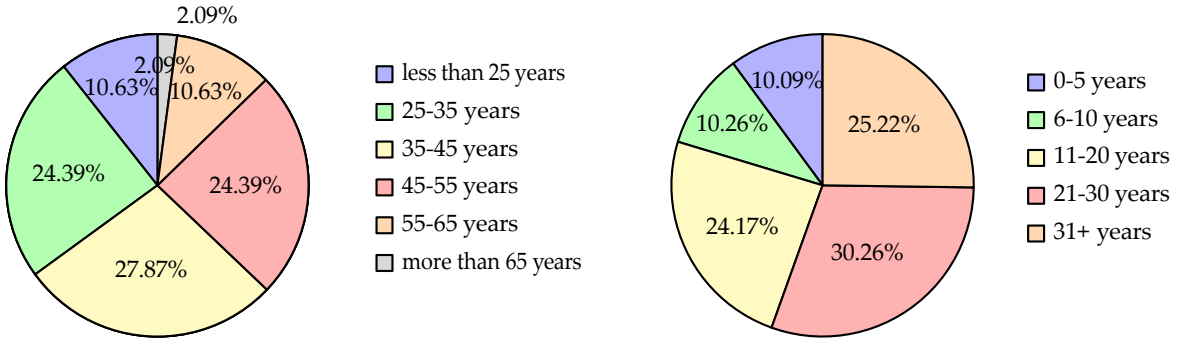


Figure 4. Age (left) and years of membership (right)

The management education of our participants is rather high, and a particular emphasis is on group, platoon, and formation leaders (Gruppen-, Zug- und Verbandsfu¨ hrer), who usually take responsibility for their units and lead operations and organizations (cf. left chart of Figure 5).

We further asked our participants to rate their perceived IT knowledge to get an idea of their skills and to be able to rate their answers in an IT-related context. Although this rating is not objective, it helps us rate comments and attitudes (cf. the right chart of Figure 5).

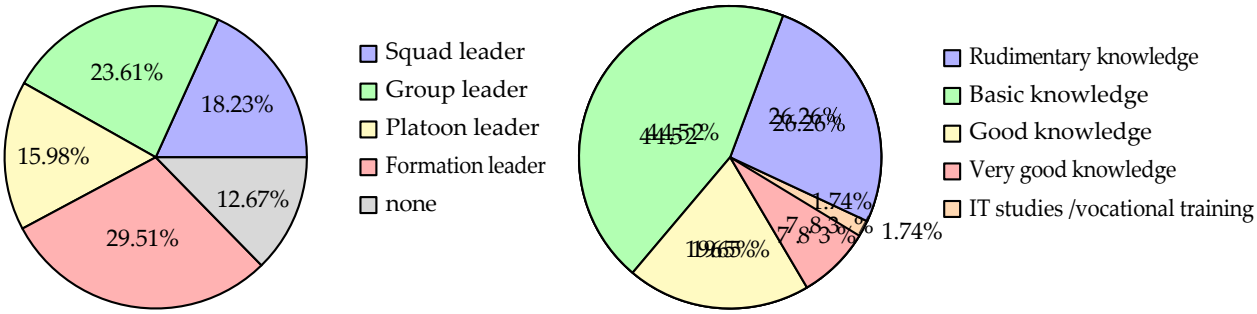


Figure 5. Level of fire department management training (left) and IT knowledge (right)

The vast majority of our participants are doing firefighting on a voluntary basis. This is not a surprise, since of the 1.385 million firefighters in Germany, 1.014 million belong to a voluntary fire brigade. Only 35 thousand firefighters are professional firefighters at governmental institutions, plus another 34 thousand firefighters work for companies. The remaining 301 thousand are members of youth fire departments. Some people also belong to both types of organizations (Deutscher Feuerwehrverband, 2021). The shares can be seen in Figure 6.

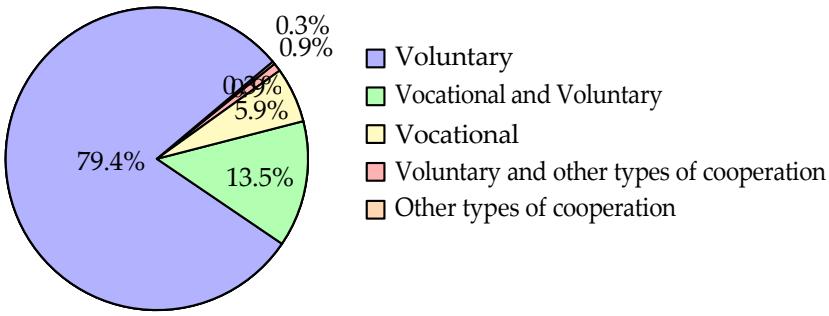


Figure 6. Types of engagement in the fire department

SURVEY EVALUATION

In the following section, we will present the knowledge-management-related aspects. We will begin with a survey of the current status of knowledge management. This includes, according to (Nonaka & Takeuchi, 1995), the acquisition, administration, and dissemination of *information* (cf. Figure 1). This is followed by questions about knowledge management and its distribution, as well as requirements from our respondents.

Information Acquisition, Administration, and Dissemination

Regarding the systems that are in use for information dissemination and the perceived availability of information, we asked the following questions:

How is new information, such as instructions (e.g., service instructions, emergency plans), processes, contact persons, etc., currently distributed in your fire department? (multiple answers possible)

The prevalent means of information management are e-mail, text messaging (such as WhatsApp and Telegram), verbal communication, and notices on a board (see Figure 7). Other platforms, such as a cloud, wiki, or forum, are barely used currently. The participants indicated the following software in the free text field: DIVERA 24/7 (10), Spond (9), Internal area of the homepage / intranet (8), APager (6), Teams (5), Alamos (4), as well as other products with less than three answers per product. The products are presented in detail in the Findings section.

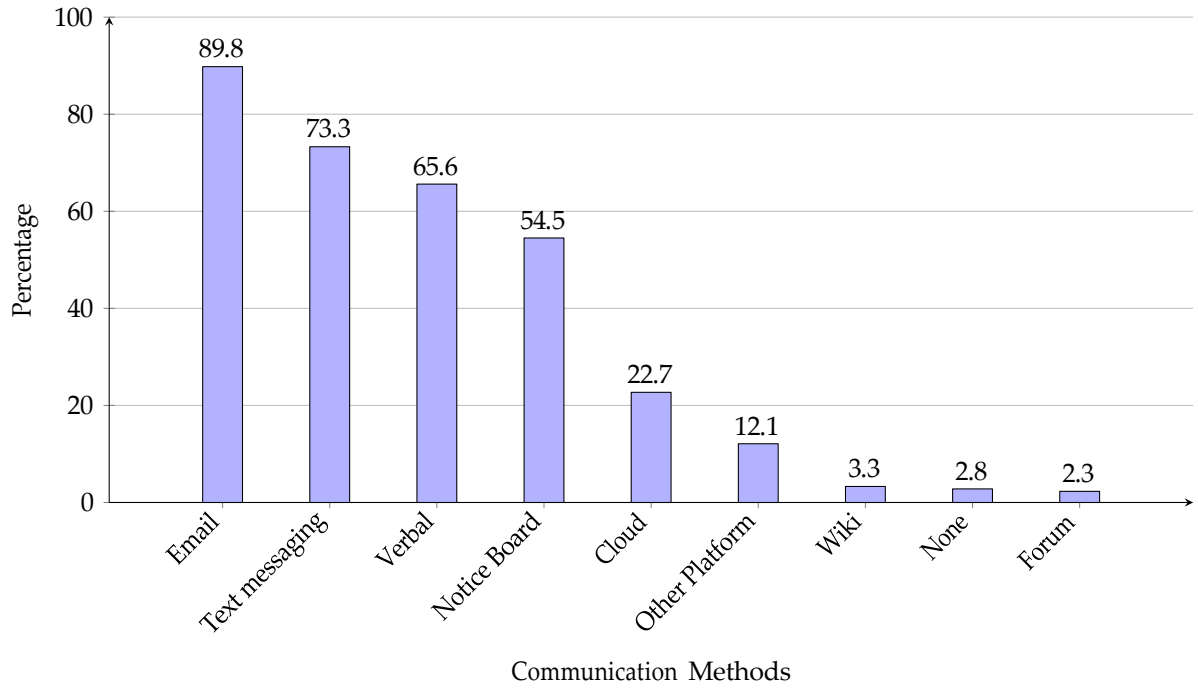


Figure 7. Distribution of communication methods in fire departments (multiple answers possible)

Rating statements towards the current situation of information and knowledge management.

We asked the following questions, the answers to which are shown in Figure 8.

- Question 1: I have the feeling that all current and new information is available to me.
- Question 2: I find it important that information and knowledge are stored in an organized way in my unit.
- Question 3: It is important to me that as much knowledge as possible is available to as many people as possible.
- Question 4: A knowledge management system should provide answers to recurring questions.
- Question 5: A knowledge management system should help to keep the content up-to-date.
- Question 6: The same information must be communicated frequently (e.g., answers to recurring questions).

Remarkably, only 53 percent of the participants stated that they agree, or rather agree, with the claim in Question 1. On the other hand, 32 percent said that they disagree, or rather disagree, with this claim. So, a significant share of the participants perceive a lack of information.

Questions 2 and 3 (cf. Figure 8) relate to the organization and dissemination of information among firefighters. The participants find it important that information is stored in an organized way and, thus, ideally, retrievable with no extra effort. Also, the transparency of information is important to keep personnel informed about everything relevant.

Questions 4 and 5 deal with the features that knowledge management should provide. Here are only two preliminary questions to get a first overview. More requirements and expectations are raised in the course of the survey.

Question 6 revealed other results than expected. Our assumption was that management personnel have to constantly answer similar questions. The response behavior is not as clear as expected but still displays the need to communicate similar information.

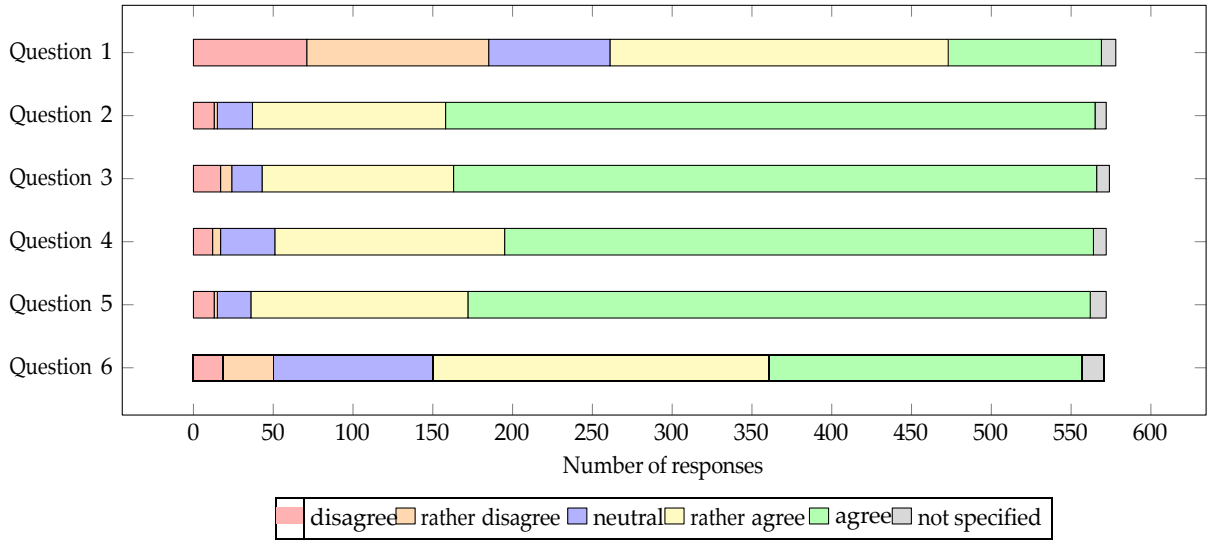


Figure 8. Statements towards information and knowledge management

Knowledge Management

Next, we surveyed the current status of knowledge management. 117 respondents (20.03 %) use a KMS, 343 (58.73 %) do not use a KMS, and 118 (20.21 %) are unsure whether their unit has a KMS in use. All respondents stating 'yes' were then asked which system they use exactly. This gives an overview of many systems that have practical applications. The whole set of software products mentioned throughout the survey is listed in Table 1.

Knowledge management systems in use

We grouped the answers with at least three mentions into the following categories:

- Information management systems / Databases: MP-Feuer (23), FeuerON (19), fireplan (8), Wasserkarte.info (3), 'database' (3) (not further specified)
- Knowledge management systems: Fireboard (19), Alamos (11), Wiki (5) (unclear about the actual usage), DIVERA 24/7 (4), L2R (3)
- Other software (depends on the usage, thus not categorised in one category above): Cloud solutions (18), Teams (3), self-development (3)

Features of knowledge management systems

We asked the participants without a KMS in their unit which features they personally would make use of. 414 would use the KMS for saving data for operations (e.g., hydrant plans or checklists), 411 would use it for providing content for basic and advanced training (e.g., training scripts), and 392 would use the system for saving instructions

for procedures (e.g., clothing exchanges or service directives). With a certain distance, 301 answered "Storing task descriptions (e.g., tasks as a manager)", 286 answered "Saving information about operations (e.g., operation reports)", and 280 answered "Managing base data (for promotions, addresses, or equipment maintenance)".

Subsequently, we asked which functions or contents should be added to or provided by an existing KMS. The answers are partially very generic; others are very specific. We tried to categorize the answers into different, throughout the process evolving, categories that match the meanings of the respondents best.

Many respondents (46) ask for a feature to provide teaching material from existing courses to either have a look at it prior to training or afterwards to refresh the knowledge of an attended course. Another 20 respondents want a KMS to help them harmonize existing teaching material to align and adjust courses' contents to have similar training for all personnel. 9 respondents request a lessons-learned approach to be able to gain knowledge from incidents or risk situations and to inform others about this issues.

Also a regular request is the storage of instructions (29) to have a central position where all applicable instructions can be found. 15 people requested an update mechanism to receive information upon changes in whatever information is stored in the KMS. Another 9 respondents want to record responsibilities and organization charts in a KMS, together with 9 answers that seek to display processes. 5 people want to have an frequently asked questions (FAQ) section, together with another 5 respondents demanding an on-boarding procedure in the KMS. 5 people requested a database of forms.

20 people requested mission-related standardized rules, emergency plans (22) and other maps relevant to an incident.

12 respondents want to store user manuals in a KMS, together with duty roster / appointments (11), equipment and material management (9), personnel administration (8), and training administration (7).

Give an example in which the knowledge management system has contributed or could contribute to solving a problem within the fire department

One respondent reports the necessity of training about 86 fire brigades and their members in their rural district according to a state-wide regulation for vegetation firefighting. The training has to be both theoretical and practical, which requires an enormous amount of effort. The respondent suggests a KMS that helps convey the training content by providing an eLearning course that adapts to the participant's prior knowledge. The course concludes with a test that certifies the theoretical knowledge, which is then applied practically in a second part. Updated material on this topic can be posted in a KMS for former attendees of the course. Likewise, 28 other respondents stated their wish to have standardized implementation of training in a KMS that adapts to the respective management level. The possibility of recovering previously attended courses is demanded, too. Besides, respondents required the standardization of knowledge bases (11) and processes (6), as well as maintaining a set of FAQs (6).

The process of onboarding new personnel is important for 14 respondents who wish to have a place where they can both store information for the people introducing new colleagues to their job as well as a possibility for new people to read up on things about their new place of duty. 4 respondents want to have a place for descriptions of their tasks to facilitate the handover of responsibilities if people in positions change.

11 respondents want to use a KMS to organize their material and knowledge, which was substantiated by others in terms of storing instructions in a sorted way (15), saving information about locations of fire hydrants and other building-related firefighting plans (14) as well as standard operating rules (9). 6 participants requested a facility to store the latest versions of forms in one central place. 13 answers indicated that an important use case would be the management of shift plans and other means to plan events with personnel. 7 people demand a way of centrally informing and being informed of news and updates.

A respondent emphasized the central aspect of a KMS that helps the users avoid unnecessary work. The respondent described that they had to maintain their data every time the program was updated and suggested a central programmer who would take care of this issue instead of dealing with it in every single unit.

Information sharing among different legal bodies

We asked the participants whether they would make use of a KMS that is run by a fire brigade from another municipality. About 83 % answered yes, while only 13 % would not use such a system. 4 % did not answer this question. As mentioned by Duscher, there is no need to compete between different fire brigades (Duscher, 2010), which might be a reason why the exact same percentage of respondents would be in favor of opening their system for members of another fire brigade (83 %), while about 13 % would not, and 4 % did not answer this question.

Sharing of information and knowledge in the line of command

Next, we wanted to know if and how our participants make their knowledge available for others (480), issue instructions (326), and forward information in the line of command (482). We were also interested in the means they use for this, which is plotted in Figure 9. It is apparent that email is the predominant system in all three scenarios of knowledge sharing, issuing instructions, and forwarding information. Verbal communication and text messaging via WhatsApp and Co. are in second and third place, respectively. The notice board is still used by half of our respondents who issue instructions. More digitally advanced systems like clouds, wikis, and forum systems are barely used. The means are complemented by the free text answers where the user referred to Teams (7), Divera 24/7 (7), eLearning like L2R (3), and paper-based notes (3).

We also asked the respondents which digital systems they would like to use to provide knowledge, which is depicted in Figure 9 in the color red. Here, emails are the prevalent means as well, followed by text messaging. The share of people who want to use a cloud is notably higher (58 %) than the current usage of cloud platforms, and the same holds for wikis (30 %) and forums (20 %) as well. We deliberately asked only for digital means of information sharing, which is why verbal and notice boards do not have a value in red. Many free text answers support this hypothesis since only two answers stated that they would like to use analogous means like a notice board.

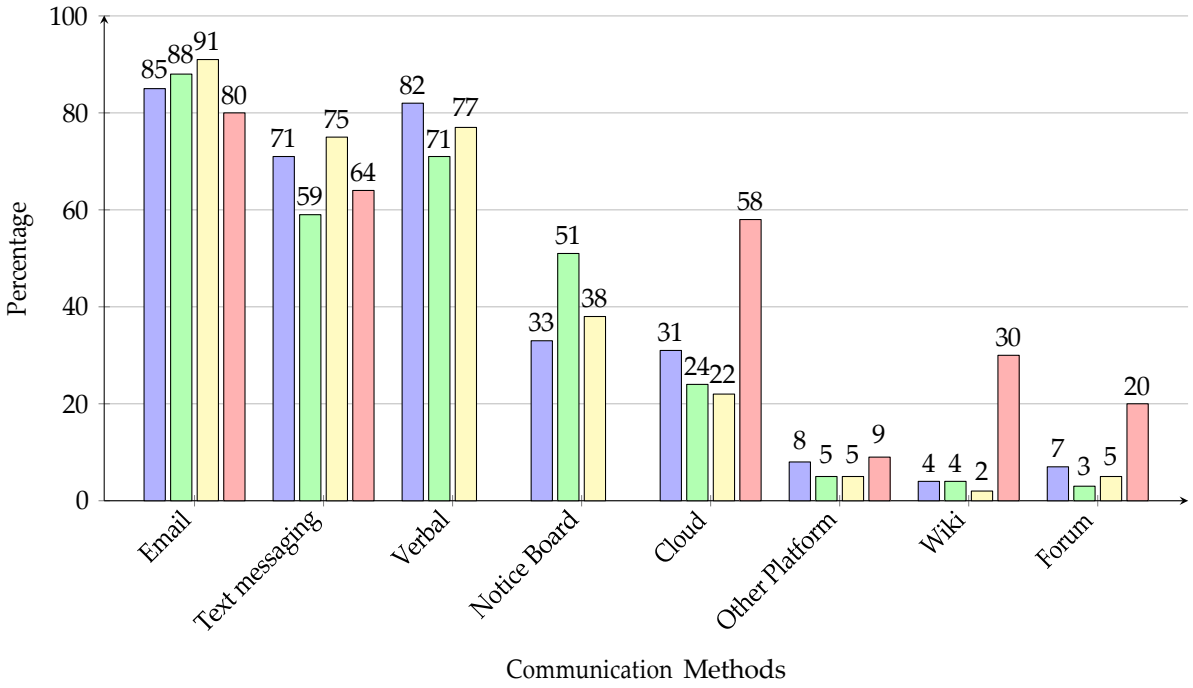


Figure 9. Distribution of means in use for knowledge sharing (blue), instruction issuing (green), information forwarding (yellow), and digital means respondents would like to use (red) (multiple answers possible)

Maintenance of knowledge in knowledge management systems

In the last section of the survey, we covered the aspect of maintaining the KMS. Therefore, we asked the respondents the following six different questions, for which the answers are shown in Figure 10.

- Question 1: It is important to me to be able to access older versions of entries.
- Question 2: It is important to me to be able to integrate / reference / transfer content from external systems.
- Question 3: It is important to me to be able to integrate content from the knowledge management system into other systems.
- Question 4: I would enter information into a knowledge management system myself.
- Question 5: I would have time to keep the knowledge I have entered up-to-date.
- Question 6: The knowledge, information, and instructions I provide are also taken into account in the long term.

Question 1 refers to the possibility for the user to access older versions of a record to keep track of changes and be able to understand its evolution and current status. Questions 2 and 3 deal with the integration of information items from external systems into the unit's KMS as well as referring to information items from other systems to avoid redundancies. Questions 4 and 5 cover the personal perception of both entering information into a KMS and keeping the provided information up-to-date. Question 6 is a more estimating question that refers to information and instruction in general that are issued and distributed by our respondents.

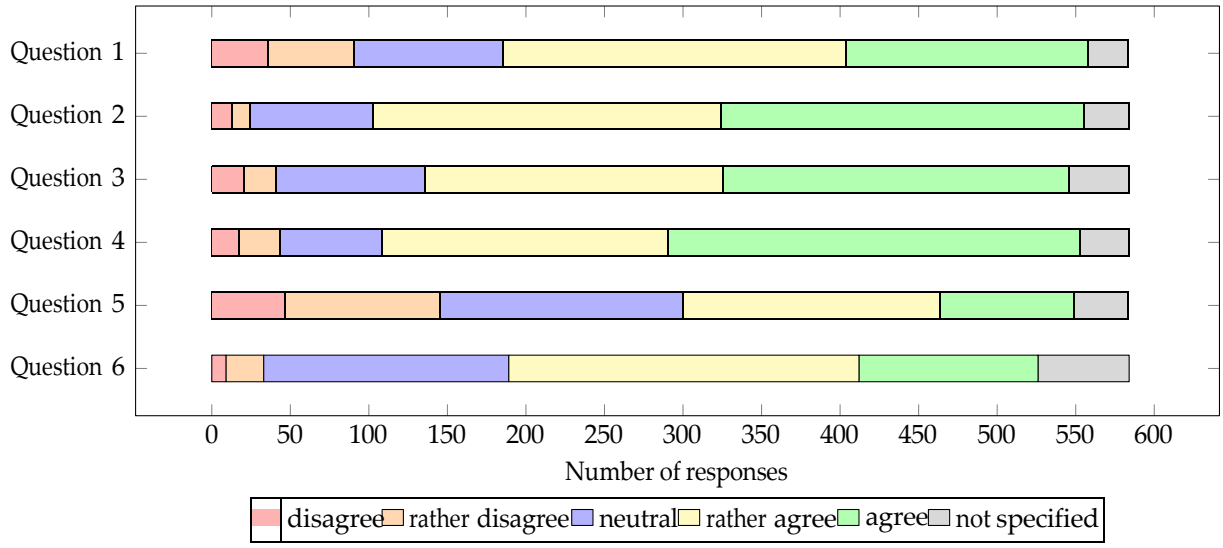


Figure 10. Statements towards information and knowledge maintenance

Final remarks on the topic of knowledge management in fire brigades in general

As the closing question, we asked the participants to write whatever they think they need to communicate to us regarding the survey's realm, which has not been addressed yet.

The most mentioned aspect is the need and requirement to have a KMS that is not only managed and maintained by a central authority (mostly state-wide) but also filled with general content (15). This is supported by the claims that the necessary efforts to run a KMS are quite high (6), as well as interfaces to other systems that are needed (4). 5 respondents explicitly stated user friendliness as an important point.

14 respondents thanked for the survey and expressed their opinion that this topic is important in modern firefighting. 4 also deliberately said that digitization in general is an important topic for the fire brigades, while 3 respondents stressed the need to have a positive balance between effort and effect. 3 respondents stated that current and existing solutions are not used enough, although they exist.

10 respondents reported that the introduction of a KMS in their unit may also be hampered by their current leadership. Some of them might object to a KMS (as something not needed) or refuse to use one because sharing information could endanger their high position (if knowledge is shared more widely). Others might only be reluctant to use digital systems in an environment that has a long tradition and well-established processes. Respondents are also rejecting a KMS because of the high demand to keep it up-to-date (5) and the costs it imposes (5). 3 respondents stated that there is no willingness to share information because others might have unwanted advantages of one's information.

Validity

The results of the survey represent the attitude of the participants towards knowledge management in fire brigades. The participants' distribution is spread across all ages and leadership levels with a tendency towards higher educated leadership levels. Since the call to participate in the survey was only distributed through digital means, this might have impacted our results as we were not able to reach people who do not use any of these means. But on the other hand, these technologies are used by a vast majority of people in Germany (Pokorny & Roose, 2020), and our intention was not to have a representative survey with exact values but to have a first estimation of this extensive topic.

FINDINGS

Although we provided a short description of what we deem an *information* management system and what a *knowledge* management system, many responses expected a KMS to be an all-in-one solution of IMS (master data about persons, equipment, and histories of trainings or incidents), a command support system (used during an emergency to provide plans, files, and procedures), and an actual KMS (dealing with training curricula, maintaining instruction and procedures, lessons learned, and best practices). Especially in the training sector, many respondents seek a system that supports their knowledge management (acquisition, administration, and dissemination).

The answers often focus on incident-related topics and the management of crises like fire hydrant maps, means to store evacuation plans, and standard operating rules for incidents. In addition, the inclusion of non-incident-related topics such as training or the management of administrative knowledge and processes indicates a more holistic view.

As seen in the red bars in Figure 9, the willingness to use digital solutions in the firefighting domain is existent. The users know these systems from other applications and want to transfer them to the fire brigades as well.

It is noteworthy that most respondents want to enter their knowledge into a KMS at first but then do not think that they have enough time to keep it up-to-date (cf. Questions 4 and 5 in Figure 10).

Next, we give a short overview of the most frequently mentioned software products throughout our survey to have a single place where we can relate to them during the presentation of the results. A summary of the software products that have been mentioned throughout our survey and are presented below can be found in Table 1.

- DIVERA 24/7[4] represents a web-based database solution primarily aimed at enhancing efficiency and precision in personnel deployment within the civil protection and disaster relief sectors. This system records, analyzes, and visually presents the availability of emergency service personnel, including firefighters and rescue services, thereby facilitating more effective resource management in emergency situations.
- Spond[5] is a digital solution focused on the administrative management of teams, primarily in sporting environments. Its functionality includes the creation of events and facilitates the exchange of a variety of data forms, such as text, images, and files, within a team setting. It is used to organize groups of several people.
- APager PRO[6], developed by Alamos GmbH[7], represents a mobile solution engineered for the enhancement of emergency alerting, response feedback, and availability planning mechanisms. Utilizing established Google and Apple transmission channels, the system ensures prompt and secure dissemination of alerts and informational messages through push notification technology. This capability is maintained even in scenarios of limited or no internet connectivity. The application further incorporates advanced geospatial features, allowing for precise display and navigation to incident sites. In addition to APager PRO, Alamos GmbH has other software and hardware solutions for alarm systems, visualization tools, and deployment tablets.
- Microsoft Teams[8] is an office application for organizational use, providing a workspace for real-time collaboration and communication. This platform integrates functionalities for meetings, file and app sharing, and allows for a more engaged and effective team environment.
- Fireboard[9], a comprehensive incident command software, is engineered to streamline the management of emergency responses. Integrating a suite of functionalities, it facilitates detailed documentation, resource allocation, and real-time situational analysis. Its modular design allows customization for specific operational needs, encompassing incident command, mapping, communication, and patient management.
- Wasserkarte.info[10] serves as a software tool for maintenance, documentation, and planning. The application offers various features like maintenance planning and automatic data transfer, enhancing the operational efficiency of water supply management.
- L2R[11] is an eLearning platform tailored to the specific needs of personnel in emergency services, including rescue, emergency medicine, and firefighting. Courses can be purchased as well as entered individually.

[4] https://www.divera247.com/ retrieved December 4, 2023.
[5] https://www.spond.com/ retrieved December 4, 2023.
[6] https://alamos.gmbh/loesungen/zusatzalarmierungen retrieved December 4, 2023.
[7] https://alamos.gmbh/ retrieved December 4, 2023.
[8] https://www.microsoft.com/en/microsoft-teams/group-chat-software retrieved December 4, 2023.
[9] https://fireboard.net/ retrieved December 4, 2023.
[10] https://www.wasserkarte.info retrieved December 4, 2023.
[11] https://www.l2r.de/ retrieved December 4, 2023.

- MP-Feuer[12] is an administration software for the fire department, rescue service, and preventive fire protection in Germany. It is established in Germany and has over 7,000 customers. MP-Fire consolidates relevant information into one, easy-to-navigate system, enabling rapid deployment across multiple operational areas.
- FeuerON[13] is a web-based software solution designed for the streamlined and transparent management of critical fire department data. It facilitates the organization of personnel, equipment, reports / incidents, and training registrations. The system offers comprehensive access to relevant data from any location with internet connectivity, enhancing operational efficiency in various areas such as personal data management, equipment and material tracking, report generation, and appointment monitoring. It features an individual rights management system, allowing administrators to control access based on user relevance.
- Fireplan[14] streamlines emergency response through its auto-dispatch and object information features, providing crucial data en route to incidents. It enhances incident management with capabilities like automatic documentation, integrated review and approval processes, and intelligent billing. The software also includes comprehensive modules for reporting, personnel, and equipment management, supporting efficient organization and analysis. Additionally, Fireplan integrates various management aspects such as training, awards, maintenance logs, and scheduling, all within a user-friendly platform.
- MediaWiki[15] is the most prominent example of a wiki system. It is the technical basis for wikis and is therefore an indispensable tool for the community-based collection and dissemination of knowledge. MediaWiki is known for its customizability and is supported by a variety of extensions and skins. It is used in a variety of projects, from large knowledge bases such as Wikipedia to small community initiatives, and features robust syntax, comprehensive user management, and detailed revision histories. MediaWiki is open source, while other wiki systems might be commercial.

CONCLUSION

We have conducted a survey with 577 responses from firefighters from all over Germany and were able to extract new findings from the data, which we will then use to elaborate on the requirements for a KMS. We can conclude that our initial hypothesis, which postulated that the spread and use of *knowledge* management systems in German fire departments is currently limited and confined to a few projects, is true. There are only very few systems in use that can be considered a KMS. This also highly depends on the individual usage of a system like a cloud or Microsoft Teams. *Information* management systems are widely used by fire brigades, as indicated by our respondents, and some examples of these systems are listed above. This survey also revealed aspects that have to be considered when designing and developing a KMS for fire brigades. The need for a system that makes use of an information sharing infrastructure is supported by the answers from our respondents, who seek an easy-to-use system where information is added only once by a superior authority and then all subsequent units can import this information items. Nevertheless, local adjustments and characteristics have to be considered since not everything can be described globally. An integrated solution that also enables the import of information from external systems would be useful to keep the maintenance of the KMS simple and automated. In addition to the insights mentioned above, our survey underscores the importance of a KMS that is applicable in both routine and exceptional operational scenarios. Our findings indicate that firefighters' familiarity with the KMS during regular duties and activities contributes to their ability to use the system effectively, even in stressful situations.

When taking the next steps in designing, developing, and implementing a KMS for emergency organisations, there are many more aspects to be taken into account. One is the rejection by end users who are not familiar with IT or deem it to be unhelpful in their daily tasks. Especially IT systems in use during the response to an emergency pose high requirements to their developers and multiple reasons for rejection (Elmasllari, 2018; Kapalo et al., 2019). Therefore, we focus at the beginning on non-incident scenarios where a KMS can have large effects like time savings, information distribution, and training as well. Nevertheless, the gained knowledge should always contribute to the cirsis management which is the core task of every emergency organisation.

Another aspect that has to be taken into account is data protection, together with data safety and security. In the EU, the General Data Protection Regulation (GDPR) poses high requirements when dealing with personal data (or data that refers to a human person) that have to be respected when developing and operating a KMS (Hu¨ttelmaier et al.,

[12] https://www.mp-feuer.de/ Retrieved December 4, 2023.
[13] https://www.feueron.de/ retrieved December 4, 2023.
[14] https://www.fireplan.de/ retrieved December 4, 2023.
[15] https://www.mediawiki.org/ retrieved December 4, 2023.

Table 1. Combined overview of software categories and functions

Software	Licence	Information Management	Knowledge Management	Other & Collaboration	Communication & Response	Alerting & Management	Administration & Training	Education
DIVERA 24/7	proprietary, 1 unit with up to 50 users free, more need to pay		✓		✓	✓		
Spond	proprietary, free to use			✓	✓			
APager	proprietary, different licenses			✓	✓	✓		
Teams	proprietary, free for non-profit organizations			✓	✓			
Alamos	proprietary, different licenses		✓			✓		
Fireboard	proprietary, different licenses		✓			✓	✓	
Wasserkarte.info	proprietary, free in basic use, more features need to pay	✓					✓	
L2R	proprietary, different licenses		✓					✓
MP-Feuer	proprietary, different licenses	✓					✓	
FeuerON	proprietary	✓					✓	
fireplan	proprietary, different licenses	✓				✓	✓	
MediaWiki	open-source	✓	✓	✓			✓	✓

CoRe Paper – Impact-driven Emergency and Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

2021). But also, the security aspect has to be taken into account (Walker, 2012) since many of the data stored is confidential, leading to the need for certain restriction policies.

A broader literature review on KMS in disaster management was performed by Dorasamy et al. Their results state that the terminology related to emergency and disaster management should be standardized to ensure the comparability of different approaches. Also, the differences and similarities between KMS and emergency management information systems should be elaborated upon (Dorasamy et al., 2013). Using standardized procedures as well as interfaces between KMS and third-party systems is an important factor when it comes to implementation. Many fire brigades already use one or another software tool, and interoperability should be possible. As indicated by Wesendrup et al., the six challenges for data management for firefighting in Germany are "cartographic data harmonization, information system (IS) standardization, information gathering from unstructured data, canonical bodies of knowledge, and data-driven firefighting support" (Wesendrup et al., 2019).

Open-source software itself is not required by many respondents in our survey, maybe due to their lack of IT knowledge. But having software that is free to use is important if it is supposed to be used by many, especially smaller fire brigades who cannot afford personnel to run the system for them (Timm et al., 2013).

All together, knowledge management offers great possibilities for modern firefighting to make use of the vast amount of knowledge and high technology. It can speed up actions, keep people informed as well as motivated, and help organizations in transition phases while preventing a loss of knowledge. To be successful, many requirements by end users have to be considered, and satisfaction and acceptance are key success factors for the helpful development and implementation of a KMS.

REFERENCES

Anand, A., Buhagiar, K., Kozachenko, E., & Parameswar, N. (2022). Exploring the role of knowledge management in contexts of crisis: A synthesis and way forward. *International Journal of Organizational Analysis*. https://doi.org/10.1108/IJOA-02-2022-3156

Bundeszentrale fu¨r politische Bildung. (2023). Bevo¨lkerung nach bundesla¨ndern - zahlen und fakten.

Deutscher Feuerwehrverband. (2017). Informationen zu altersgrenzen in der jugendfeuerwehr, freiwilligen feuerwehr, berufsfeuerwehr und werkfeuerwehr.

Deutscher Feuerwehrverband. (2021). Mitgliederstatistik der deutschen feuerwehren.

Deutsches Institut fu¨r Normung. (2022). Din iso 30401:2022-11, wissensmanagementsysteme - anforderungen (ISO 30401:2018 amd.1:2022). https://doi.org/10.31030/3384285

Dorasamy, M., Raman, M., & Kaliannan, M. (2013). Knowledge management systems in support of disasters management: A two decade review. *Technological Forecasting and Social Change*, *80*(9), 1834–1853. https://doi.org/10.1016/j.techfore.2012.12.008

Duscher, S. (2010). *Die auswirkungen von motivation auf die wissensweitergabe in einer freiwilligen feuerwehr* [Diplomarbeit]. FHWien der WKW. Wien.

Elmasllari, E. (2018). Why it systems for emergency response get rejected: Examining responders' attitude to it. *ISCRAM*.

Hu¨ttelmaier, T., Weber, B., & Harder, Y. (2021). Die gefahrenabwehr im zeichen der digitalisierung. *Kohlhammer*, *2021*.

Kapalo, K. A., Wisniewski, P., & LaViola Jr., J. J. (2019). First in, left out : Current technological limitations from the perspective of fire engine companies. *ISCRAM 2019*.

Mozgova, I., Koepler, O., Kraft, A., Lachmayer, R., & Auer, S. (2020). Research data management system for a large collaborative project. *Balancing Innovation and operation*. https://doi.org/10.35199/norddesign2020.48

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company*. Oxford University Press.

North, K. (2016). Die wissenstreppe. In *Wissensorientierte unternehmensfu¨hrung: Wissensmanagement gestalten* (pp. 33–65). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-11643-9_3

Oktari, R., Munadi, K., Idroes, R., & Sofyan, H. (2020). Knowledge management practices in disaster management: Systematic review. *International Journal of Disaster Risk Reduction*, *51*. https://doi.org/10.1016/j.ijdrr.2020.101881

Ortmanns, W., & Sonntag, R. (2023). *Umfragen erstellen und auswerten*. UVK Verlag. https://doi.org/10.24053/9783739882413

Pickles, A. (2004). A knowledge management infrastructure for the nsw fire brigades. *Australian Journal of Emergency Management, The*, *19*(2), 21–24.

Pokorny, S., & Roose, J. (2020). Die eignung von umfragemethoden.

Sauerland, T., Werner, T., Rammert, S., Habig, T., Marterer, R., Schultz, A., & Scheuermann, L. (2020). Digitale transformationin der zivilen gefahrenabwehr. *Studie im Auftrag der Vereinigung zur Fo¨rderung des deutschen Brandschutzes e.V.*

Timm, E., Alsbach, S., & Wimmer, M. A. (2013). Wissens- und Kollaborationsplattform im Brand- und Katastrophenschutz: Open-Source als kostengu¨nstige Option? In M. Horbach (Ed.), *Informatik 2013 – informatik angepasst an mensch, organisation und umwelt* (pp. 1634–1648). Gesellschaft fu¨r Informatik e.V.

Walker, J. (2012). Cyber security concerns for emergency management. *Emergency Management; Eksioglu, B., Ed.; InTech: Rijeka, Croatia*, 39–59.

Wesendrup, K., Rupp, N., Widera, A., & Hellingrath, B. (2019). Challenges and Trends of Data Management for Firefighting in Germany and the Netherlands. *ISCRAM 2019*.

Ziehr, C., Skorupa, F., Schmitt, K., & Kopp, M. (2017). Berliner feuerwehr: Wiki-systeme fu¨r den einsatz- und ru¨ckwa¨rtigen dienst. *Brandschutz, W. Kohlhammer, Stuttgart*, (5/2017), 342 ff.

Cascading supply chain disruption: impact on water industry during COVID-19

Lijo John
Birla Institute of Technology Pilani-Dubai Campus
Hanken School of Economics, Helsinki
lijo@dubai.bits-pilani.ac.in
lijo.john@hanken.fi

Wojciech Piotrowicz
Hanken School of Economics, Helsinki
wojciech.piotrowicz@hanken.fi

Aino Ruggiero
Hanken School of Economics, Helsinki
aino.ruggiero@hanken.fi

ABSTRACT

The impact of COVID-19 on the lives of people and businesses across the globe was devastating. While governments across the world had undertaken a slew of measures to control the spread of the COVID-19 virus within their geography, many of these measures had long and unintended consequences. The restrictions imposed by the governments on the movement of people and goods across the world brought supply chains to a grinding halt. This study identifies the cascading effects of supply chain disruptions on the water sector. Since these systems are closely integrated and the impact of COVID-19 needs to be analysed at a much broader level, this study uses a systems-thinking approach to study the effect of supply chain disruptions on water services. The study develops a causal loop model to gain further insight into how supply chain disruptions caused by COVID-19 affected the coping capabilities of society and how critical services were affected.

Keywords

Qualitative systems model, COVID-19, water sector, supply chain disruptions, disaster management

INTRODUCTION

Supply chain disruption (SCD) is an unexpected event that stops or slows the normal flow of material with potentially negative consequences to supply chain members (Chopra and Sodhi, 2004). At a broad level, SCDs are usually classified based on their causes, such as acts of nature (e.g. flooding, earthquakes, hurricanes and pandemics). While this cause-based classification identifies the underlying reasons for these disruptions, it is often much more useful to gain further insights into the effects and stage of disruption and this has led to other classifications of SCDs. For instance, Ivanov et al. (2017) classified SCD based on the level or echelon at which the disruption has occurred. Chopra and Sodhi (2014) classified SCD based on the causes other than n atural or manmade disasters and identified that SCDs can be caused by delays in commerce systems, forecasts, intellectual property, etc. Other classifications for SCDs are based on frequency of occurrence (Tang et al., 2014), nature and their source of origin (Christopher et al., 2011) and who they affect, from broad to specific (Dolgui and Ivanov, 2021). The impact of SCDs on businesses can be operational, marketing and financial (Katsaliaki et al., 2021). Operational impacts can be product unavailability, partially fulfilled orders, late deliveries, logistic challenges, etc. (Jabbarzadeh et al., 2018; Wagner and Neshat, 2012). While marketing impact can be increase in customer

WiP Paper – Impact-driven Emergency and Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

complaints, damage to brand reputation, loss of customers, breach of supplier contracts, penalties etc. (Ponomarov and Holcomb, 2009). And finally, financial impacts are, loss of sales and revenue, reduced market share, complaints, damage to brand reputation, loss of customers, breach of supplier contracts, penalties etc. (Ponomarov and Holcomb, 2009).

However, what has been lacking in studies related to SCDs is their impact on the security of the supply of critical services, such as quality of water for daily use including drinking and bathing. The revised Drinking water directive (2020) by the EU parliament defines essential quality standards for water intended for human consumption. It requires Member States to regularly monitor the quality of water by using a 'sampling points' method. The revised directive updates existing safety standards and improves access to safe drinking water along the lines of the latest recommendations of the World Health Organization (Curmei and Kurrer, 2023). Furthermore, the directive also focuses on building more resilient drinking water system and designates the water infrastructure as a "critical entity" (European Commission, 2020). This has led to the creation of a waster security plan for prompt response and recovery from intentional or unintentional contaminations affecting the drinking water supply system, reducing the impact severity such as the potential escalation to service disruption and disease outbreaks which could have catastrophic consequences (Teixeira et al. 2019).

During the peak COVID-19, the focus on personal hygiene and cleanliness has been at the peak. This had led to rapid increase in the domestic water consumption. However, the lockdown also had a negative impact on the businesses, and this led to a rapid drop in the industrial water consumption. Overall, this mismatch in the demand profile had affected the revenues of the water companies and further affected the operations of many water companies. World bank (2020) estimated that there was a drop of about 40% where many countries had suspended the water charges to give some financial relief to their citizens during the economic slowdown during COVID-19. Furthermore, there were widespread interruption of operations due to absenteeism, labor cost and procurement of adequate personal protective equipment for staffs and limited supply of essential chemicals for water treatment during the COVID-19 pandemic were other challenges COVID-19 posed to the water sector (Joo and Trindade, 2023). This was further complicated with SCDs owing to the internal restriction set by multiple countries affecting both capacity and transportation of necessary chemicals and spare parts used by the water management companies to ensure safe water services. Thus, in this study, we explore the impact of the SCD caused by COVID-19 on the water industry and how it can contribute to the conversation of security of supply (water) of daily operations of a water utility's system and help to detect other safety issues as well as supply chain disruptions (e.g. shortage of key workers; shortage of key chemical supplies).

BRIEF LITERATURE REVIEW

Supply Chain Disruption

In decades, supply chains have moved towards efficiency and cost reduction. To improve the efficiency, economies of scale, reduction in safety stock and focus on the sole source of suppliers has been leveraged by many supply chain managers (Parmigiani et al., 2011). However, the major side effect of these activities was that the supply chains became extremely fragile and would break down even at the slightest disruption since it lacked enough cushion (Ivanov et al., 2016). This was most evident during the COVID-19 pandemic, when supply chains were under severe stress owing to supply and demand shocks (Atkinson et al., 2020).

SCD severity is influenced by design characteristics and mitigation capabilities, highlighting the need for resilience planning and business continuity planning (Craighead et al., 2007). Furthermore, it is seldom that SCD effects remain localized and often they propagate along the supply chain. They severity and intensity of SCD propagation depends on the nature of disruption, structure of supply chains, and managerial decision-making (Scheibe and Blackhurst, 2018). SCDs often has a cascading effect across the entire supply chain network (SCN). The propagation of a disruption across the SCN and its associated impact is called the ripple effect and is distinct from the well-known bullwhip effect (Dolgui et al., 2018). In more severe cases, the ripple effect can cause some nodes and arcs in the SCN to become temporarily dysfunctional (e.g. due to a material shortage and capacity loss) (Dolgui et al., 2018).

To understand the supply chain ripple effect, it is imperative to understand risk propagation in the SCN. In practice, a disruption to an SCN often begins locally, with a trigger event affecting any one or a group of nodes in the network usually located within the same vicinity. This impact spreads to other firms (or nodes) through their relationships. Consequently, the impact might propagate to parts of the network through a set of cascading effects and affect nodes that are far from the origin of the impact (Scala et al., 2021). The extent to which the

impact propagates in the supply network is also a testament to the resilience of the network (Li and Zobel, 2020). Dolgui et al. (2018) discussed the main reasons for the systemic vulnerabilities that make SCNs more prone to the ripple effect when presented with a disruption in any or multiple nodes within the SCN. These reasons can be broadly classified into four main categories: sourcing strategy risks, production planning risks, inventory management risk and control risk.

COVID-19 and Water Sector

Prior to COVID-19, the water industry across the globe were impacted by five major trends. These included, (a) global warming led increase in extreme floods and droughts, challenging the resilience of water and sanitation systems, (b) increasing number of people living in areas facing water stress (currently 2 billion), which increases water supply vulnerabilities, (c) rapid urbaniza tion, which strains existing water resources and ecosystems, (d) the emergence of megacities, which adds the challenge of extending water and sanitation services to about 1 billio n people living in informal settlements not served by water grids, (e) aging infrastructure, which has increased pressure to accelerate investments in more advanced markets, following decades of underinvestment (Butler et al., 2020). Along with the pre-existing challenges, some of the major impact of COVID-19 on water services are summarized below.

With the nationwide lockdowns, there was a marked increase in the household consumption of water and a significant decrease in the non-household or industrial consumption. With the increase in the household consumption, thereby increasing the cost of water production (Ong and Nielsen, 2020). However, revenue from metered household customers has increased because of increased demand and companies with comparatively more metered customers saw a greater increase in household revenue in the short -term, while the ones with industrial consumers saw a significant drop in their revenues (Ong and Nielsen, 2020).

On the other hand, the Health and Safety Executive (HSE) in the UK advised business including water sector to prioritize becoming "COVID-secure" by adapting to current guidance and implementing measures to control the risk of COVID-19 to protect workers and others. These steps included carrying out COVID-19 risk assessments adhering to HSE guidance, enhancing cleaning, hand washing, and hygiene procedures, taking all reasonable steps to help people work from home, etc. To comply with these mandates, water companies have had to adapt various measures, including regulatory inspections, limited site visits, phone calls (remote working), and collection of visual evidence (photos and video footage), etc. (Frontier, 2020; HSE, 2020). The pandemic had disrupted the water sector's revenue and growth projections. Hence there was a need for readjustment for the rest of the Asset Management Period (AMP) to reprioritize projects supported by Ofwat (England and Wales agency that Promotes competition, sets price limits, ensures that water companies can finance and carry out their functions, promotes economy and efficiency) (Horton and Laikin, 2020). This further affected the capital investments both in term so of maintenance and upgrading the water systems. The cost structures and lack of funding sources in the post pandemic environment put halt on the new capital expenditures for both maintenance and modernization of the water system. The tariffs were insufficient to recover operating costs and the water utilities require support from other sources, usually the government budget had also reduced during peak COVID-19 (Butler, et al., 2020). A decline in demand from large industrial and commercial users due to lockdowns and travel restrictions significantly reduced revenues to water utilities. Specific measures adopted by the governments like (a) deferments on or exemptions from utility bill payments for vulnerable groups, (b) moratoriums on cutting off the water supply, and (c) suspensions of meter reading and invoicing also a ffected the revenues of the water industry (Butler, et al., 2020).

WASH and COVID-19

The water supply and sanitation services must be considered within a larger framework of integrated water management, water quality, ecosystem health and governance and is crucial for public health and human development (Gaddis et al., 2019). Mukhtarov et al., (2022) argued that COVID-19 and water management will have long term mutual impact on each other. The direct impact includes prioritizing WASH at the expense of other crucial challenges such as climate change adaptation, land use patterns, ecosystem h ealth and energy. Furthermore, lesser public funds will be made available for water infrastructure and management and the danger of oversized influence of private financial capital in the water sector might cause lower quality of water related infrastructure. On the other hand, the long-term indirect impact will be securitisation of WASH that may lead to day-to-day and fragmented management. With private financial take-over of water infrastructure and services, the attention to techno-fixes with a relative neglect of social and political aspects of water management and governance may be affected in long term and might as well take a back seat in long run .

SYSTEMS THINKING AND METHODOLOGICAL FRAMEWORK

Systems thinking has been used as a tool to study complex interactions in supply chain research. The "system" is defined as 'a whole consisting of two or more parts, where each of which can affect the performance or properties of the whole, none of which can have an independent effect overall, and no subgroup of which can have an independent effect on the "whole". In brief, a system is a whole that cannot be divided into independent parts or subgroups of parts' (Ackoff, 2010). In the same vein, Ackoff (2010) defined systems thinking as looking at relationships (rather than unrelated objects), connectedness, process (rather than structure), the whole (rather than just its parts), the patterns (rather than the contents) of a system and context (Ackoff, 2010). This view of systems thinking aligns with the supply chain function, in which various aspects of the supply chain need to relate to an established process to communicate with various parts of the supply chain to fulfil consumers' demands (Wilden et al., 2022). According to Wieland et al. (2023), supply chains are not only complex adaptive systems but a lso have close connections with social, economic, ecological and political contexts within which they are embedded. While not referring to systems thinking directly, Sarkis et al. (2020) took systems view to critique the weaknesses in the supply chain and promote resilience measures, in addition to detailed lessons learned from a post -COVID perspective (Sarkis, 2020). Elias et al. (2021) used systems thinking to holistically analyze the complexities involved in the adoption of a sustainable wood supply chain in Amazon and to develop strategic interventions to improve the system.

System dynamics, a research approach based on systems thinking, often uses group model building (GMB) to identify complex relationships between various entities and subsystems (Vennix et al., 1999). GMB has been the favored approach for the system dynamist for three main reasons: first, to capture the required knowledge in the mental models through the various stakeholders; second, to increase the chances of implementation of model results; and third, to enhance the stakeholder learning process (Vennix et a l., 1997). GMB exercises allow stakeholders to voice their individual perceptions of reality and help them develop a shared mental model of their perceptions as a group (Elias, 2016).

For this study, an exploratory approach using the GMB method was adopted to gain an understanding of the complex nature of SCDs due to COVID-19 and its impact on the water sector (Rouwette et al., 2002). Due to the exploratory nature of the research, the discussions were semi-structured to enable the researchers to discover the varying causal influences that the participants perceived to occur in practice (Graham et al., 199 2). The subsequent sections broadly explain the main parts of the research process. The research process was divided into four ma in steps, as described below (Hovmand et al., 2021).

- The first step included an initial literature review and identification of the main impact on the water sector due to SCDs caused by COVID-19
- The second stage involved initial data collection based on a questionnaire developed by the gatekeepers, which was shared with the sCience and human factOr for Resilient sociEty (CORE) consortium members to gather insights into the impacts of supply chain disruption on the water sector due to COVID-19.
- The third stage involved a GMB workshop in Helsinki, in which the participants developed an initial set of relationships between the variables and helped identify an initial set of relationships. As part of initiating GMB, experts in HUMLOG were contacted and invited to be part of GMB.
- The fourth stage involved final model building and validation, where the model was shared with members of the CORE consortium project.

CAUSAL LOOP MODEL FOR WATER SECTOR

The impact of COVID-19 on water sector has not received much attention. However, it is not to say that the impact has been any less. This section discusses the main impacts COVID-19 had on water sector through a causal loop model developed through GMB.

The lockdowns had a significant impact on the water industry. Due to the lockdowns and increase in the remote working, a lot more people have spent significant amount of time indoors. Therefore, the domestic water consumption had increased significantly. While, on the other hand, the industrial water consumption had reduced. this had significant impact on the revenues generated by the water companies. However, the lockdown induced domestic consumption had a lot more impact than merely an increase in the domestic water consumption. With a significant amount of time spent indoors, people started using more water at home for water intensive activities such as gardening, swimming, etc. This is represented by the RW1 is a positively reinforcing loop which captures the impact of household water consumption for WASH demand which has a positive impact on reducing the risk of infection.

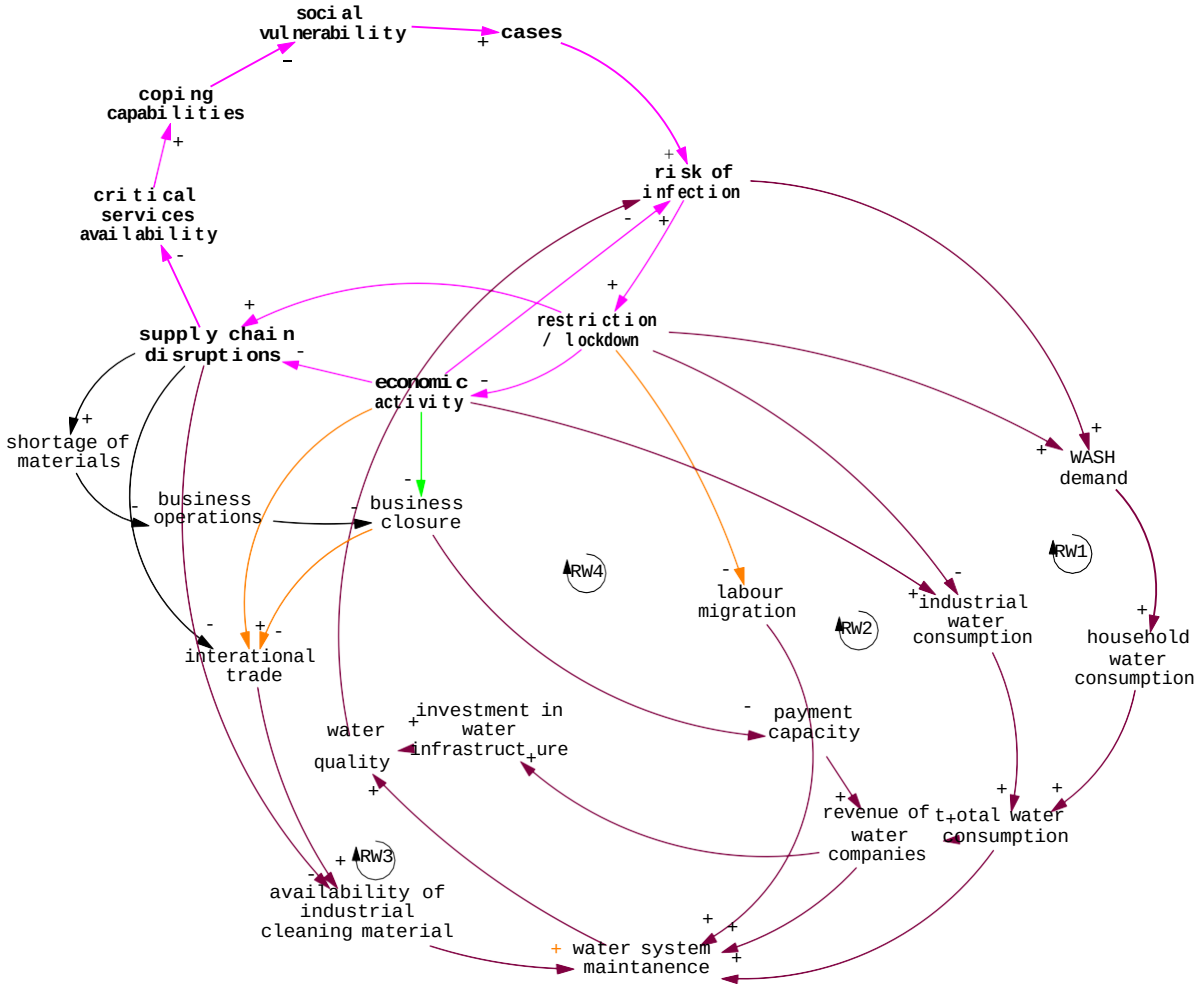


Figure 1. Causal Loop Model

However, this increase in the domestic water consumption had other unexpected impact on the water systems as well. It was observed that people were flushing down and putting things that were not supposed to be put into the drains. For instance, due to the shortages in the toilet paper (owing to international lockdowns and supply cha in disruptions), many were forced to use alternatives which were not supposed to be flushed in the drains. These kinds of behaviours had cost significant blockages along the drains and sewers, which had put additional pressure on the water companies in terms of maintenance. Thus, it became significantly more expensive for water and water systems management. Furthermore, the household demand offset the reduction in the non-household demand and as a result the water production cost for the companies had increased significantly during this time . This is represented by the negative reinforcing loops RW2 in the figure 1. RW2 captures the impact of reduction in the revenues of water companies due to that restrictions and lockdowns and thereby the substantial reduction in the industrial water consumption along with the increased expense to maintain the water systems from collapsing due to the excess stress on the already aging water filtration systems due the increase in the domestic water consumption.

The supply chain disruption due to COVID-19 had significant impact on the quality of water itself. While multiple transmission modes of COVID-19 were being closely evaluated, it was identified that COVID-19 virus could be transmitted through water as well, especially through the sewage systems from residentia l (toilet, shower, washing, etc.) and through medical centres. Thus is became the priority of the both local and national governments to ensure that water treatment was carried out properly and the water companies had been mandated to ensure that there is no lax in the water treatment protocols. However, frequent water quality checks revealed that there were increased viral levels even in the treated wastewater which indicated that the wastewater treatment was not carried out with adequate efficiency to remove COVID-19 viruses. One of the main reasons cited by the water companies for this lack of efficiency in filtration systems was due to the age-old equipment that had not received proper maintenance or in many cases were passed their due life. Secondly, it was also identified that in many cases the water companies were not able to source the chemicals and components for the filtration systems due to the supply chain disruptions. Since many are of the chemicals used in the water filtration systems are not manufactured within

Europe and is usually imported from Asia , the disruption of the international trade had led to a severe shortage of chemicals required for the water filtration. These challenges to water is represented by the reinforcing loop RW3. However, RW3 is a negatively reinforcing loop representing the impact of supply chain disruptions due to the restrictions and lockdowns on the availability of industrial cleaning material which had an impact on the water maintenance system and thereby the quality of water increasing the risk of infection through the water systems.

Finally, COVID-19 severely affected the modernization and capacity expansion plans of the water sector. Due to the reduced revenues and supply chain disruptions, many water companies decided to halt or slow down engineering investments in improving the water and sanitation systems. For instance, Ofwat, the water management authority in the UK had given clear instructions to the water companies to meet core service obligations which is, to provide water and wastewater services and not to prioritise engineering projects in the meanwhile. This was also affected due to the lack of temporary and seasonal labourers who are critical in large scale engineering projects. As mentioned earlier, the availability of these labourers was affected due to the border closures and lockdowns. The supply chain disruptions also had an impact on procuring raw materials and parts which would have been used in these engineering projects. Furthermore, the COVID-19 pandemic had further disrupted the sectors demand and growth projections. With an increase in the domestic water consumption, many were wary about the long term shift in the water consumption pattern. Another major impact on the revenues of the water companies was the non-payment of dues by the existing businesses. Due to the lockdown, both industrial and non-industrial small-scale businesses had to be shut down temporarily. This had affected the revenue streams of these small businesses and many of them did not have the capability to pay for the outstanding bills of the water companies. This this led to an increase in the accounts receivable for the water companies which further affected the revenues and financial outlay for large scale engineering projects. This is shown by the reinforcing loop RW 4 in figure 1.

As shown in figure 1, the water quality has a direct impact on the risk of infection , which further effects water for WASH. while the risk of infection leads to further restrictions or lockdown, it has a direct impact on the industrial water consumption. The reduction in the economic activities lead to business closures and the payment capability of both individuals and businesses thereby further reducing the revenue for water companies. Furthermore, the supply chain disruptions Leeds shortage of materials and availability of industrial cleaning items. All these factors affect the maintenance of water systems which further negatively affects quality of the water and the loop continues.

CONCLUSION

Global supply chains and connectivity make it possible for us to source, produce, and consume goods from all over the world through the integrated network of supply chains that run continually throughout. And thus, business have realized the importance of keeping the supply operational and spend billions to ensure that there are no disruptions to this SC operations. The effects of SCDs on enterprises have been thoroughly investigated in the past. Nevertheless, little focus has been placed on how they affect people's lives on day-to-day basis and, consequently, the vulnerability of society. Business owners and policy makers have become aware of this due to the recent global SCDs due to COVID-19. Previous studies, while having studied exogenous risks from SCDs, usually do not consider both supply-and demand-side risks caused by SCDs. This study offers a novel outlook into the role of exogenous risks from SCDs, especially in the water sector, which can have both supply and demand implications. Furthermore, the study uses a system thinking approach to develop a holistic view of how SCDs produce ripple effects on the water sector. While this study offers a holistic perspective using systems thinking, more empirical research needs to be conducted to quantify the real impact of SCDs across the water value chain.

This study was carried out as part of Work Package 4, Task 4.3, under the CORE project, to study the impact of SCDs due to COVID-19 on creating social vulnerabilities. This was explored through the impact of SCDs on the water sector. Since this sector is critical to the functioning of society, it is imperative to ensure that it functions without disruption. The study also explores how SCDs affected the operational capabilities of the water sector and how they influenced both short-term and long-term operational capabilities.

ACKNOWLEDGMENTS

This study is undertaken as a part of Work Package 4, under the CORE project (no. 101021746) funded by the European Union under grant agreement no. 101021746.

REFERENCES

Ackoff, R. L. (2010). Systems thinking for curious managers. Triarchy Press.

Atkinson, C. L., McCue, C., Prier, E., & Atkinson, A. M. (2020). Supply chain manipulation, misrepresentation, and magical thinking during the COVID-19 pandemic. The American Review of Public Administration, 50(6-7), 628-634.

Butler, G. Pilotto, R.G., Hong, Y., Mutambatsere, E. (2020), The Impact of COVID-19 on the Water and Sanitation Sector, available at: https://www.ifc.org/wps/wcm/connect/126b1a18-23d9-46f3-beb7-047c20885bf6/The+Impact+of+COVID_Water%26Sanitation_final_web.pdf?MOD=AJPERES&CVID=ncaG-hA

Chopra, S., and M. S. Sodhi. (2014). "Reducing the Risk of Supply Chain Disruptions." MIT Sloan Management Review 55 (3): 73–80.

Christopher, M., Mena, C., Khan, O., & Yurt, O. (2011). Approaches to managing global sourcing risk. Supply chain management: An international journal, 16(2), 67-81.

Curemei, M., M. and Kurrer, C. (2023). WATER PROTECTION AND MANAGEMENT, Fact Sheets on the European Union - 2023.Available at: www.europarl.europa.eu/factsheets/en

Dolgui, A., & Ivanov, D. (2021). Ripple effect and supply chain disruption management: new trends and research directions. International Journal of Production Research, 59(1), 102-109.

Dolgui, A., Ivanov, D., & Sokolov, B. (2018). Ripple effect in the supply chain: an analysis and recent literature. International Journal of Production Research, 56(1-2), 414-430

Elias, A. A. (2016). Analysing the stakes of stakeholders in research and development project management: a systems approach. R&D Management, 46(4), 749-760.

Elias, A. A., Donadelli, F., Paiva, E. L., & Bacic Araujo, P. P. (2021). Analysing the complexities of sustainable wood supply chain in the Amazon: a systems thinking approach. The International Journal of Logistics Management, 32(4), 1481-1505.

Frontier (2020) Economic Impacts of COVID-19 on the Water Sector, Available at: https://www.water.org.uk/wp-content/uploads/2020/12/Impact-of-COVID-19-on-the-water-sector_FINAL-REPORT-STC-141220.pdf

Gaddis, E., Grellier, J., Grobicki, A., Hay, R., Mirumachi, N., Mukhtarov, F., & Rast, W. (2019). Freshwater policy. In global environment outlook-GEO-6: Healthy planet, healthy people. UN Environment.

Graham, A. K., Morecroft, J. D., Senge, P. M., & Sterman, J. D. (1992). Model-supported case studies for management education. European Journal of Operational Research, 59(1), 151-166.

Health and Safety Executive (HSE) HSE Urges Businesses to Become COVID-Secure, London (2020), Available at: https://press.hse.gov.uk/2020/07/02/hse-urges-businesses-to-become-covid-secure/ ,

Horton, P., Laikin, R, AMP7 and COVID-19 - Future Water Asks What Now for the Water Sector? (2020) Available at: https://www.waterbriefing.org/home/water-issues/item/17164-future-water-asks-what-now-for-the-water-sector? ,

Hovmand, P. S., Andersen, D. F., Rouwette, E., Richardson, G. P., Rux, K., & Calhoun, A. (2012). Group model-building 'scripts' as a collaborative planning tool. Systems Research and Behavioral Science, 29(2), 179-193.

Ivanov, D. (2017). Simulation-based ripple effect modelling in the supply chain. International Journal of Production Research, 55(7), 2083-2101.

Ivanov, D. (2017). Simulation-based ripple effect modelling in the supply chain. International Journal of Production Research, 55(7), 2083-2101

Jabbarzadeh, A., Fahimnia, B., & Sabouhi, F. (2018). Resilient and sustainable supply chain design: sustainability analysis under disruption risks. International Journal of Production Research, 56(17), 5945-5968.

Joo, K. And Trindade, C. M. (2023), Opinion of the European Economic and Social Committee on 'Access to water: tackling water poverty and its implications for social policy' Official Journal of the European Union, Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52023IE1665

Katsaliaki, K., Galetsi, P., & Kumar, S. (2021). Supply chain disruptions and resilience: A major review and future research agenda. Annals of Operations Research, 319, 965–1002.

Li, Y., & Zobel, C. W. (2020). Exploring supply chain network resilience in the presence of the ripple effect.

International Journal of Production Economics, 228, 107693.

Mukhtarov, F., Papyrakis, E., & Rieger, M. (2022). Covid-19 and water. In Covid-19 and international development (pp. 157-173). Springer, Cham.

Ong, A., and Nielsen, E, (2020), Economic impacts of covid19 on the water sector. Frontier economics. Available at: https://www.water.org.uk/wp-content/uploads/2020/12/Impact-of-COVID-19-on-the-water-sector_FINAL-REPORT-STC-141220.pdf

Parmigiani, A., Klassen, R. D., & Russo, M. V. (2011). Efficiency meets accountability: Performance implications of supply chain configuration, control, and capabilities. Journal of operations management, 29(3), 212-223.

Ponomarov, S. Y., & Holcomb, M. C. (2009). Understanding the concept of supply chain resilience. The international journal of logistics management, 20(1), 124-143.

Rouwette, E. A., Vennix, J. A., & Mullekom, T. V. (2002). Group model building effectiveness: a review of assessment studies. System Dynamics Review: The Journal of the System Dynamics Society, 18(1), 5 -45.

Sarkis, J., Dewick, P., Hofstetter, J. S., & Schröder, P. (2020). Overcoming the arrogance of ignorance: supply-chain lessons from COVID-19 for climate shocks. One Earth, 3(1), 9-12.

Scala, B., & Lindsay, C. F. (2021). Supply chain resilience during pandemic disruption: evidence from healthcare. Supply Chain Management: An International Journal, 26(6), 672 -688.

Scheibe, K., & Blackhurst, J. (2018). Supply chain disruption propagation: a systemic risk and normal accident theory perspective. International Journal of Production Research, 56 (1 -2), 43 - 59.

Tang, S. Y., H. Gurnani, and D. Gupta. (2014). "Managing Disruptions in Decentralized Supply Chains with Endogenous Supply Process Reliability." Production and Operations Management 23 (7): 1198 –1211

Teixeira, R.Carmi, O.Raich, J.Gattinesi, P.Hohenblum, P. 2019. Guidance on the Production of a Water Security Plan for Drinking Water Supply. Available at: https://publications.jrc.ec.europa.eu/repository/handle/JRC116548

Vennix, J. A. (1999). Group model-building: tackling messy problems. System Dynamics Review: The Journal of the System Dynamics Society, 15(4), 379 -401.

Vennix, J. A., Andersen, D. F., & Richardson, G. P. (1997). Special issue on group model building. System Dynamics Review, 13(2).

Wagner, S. M., & Neshat, N. (2012). A comparison of supply chain vulnerability indices for different categories of firms. International Journal of Production Research, 50(11), 2877-2891.

Wieland, A., Stevenson, M., Melnyk, S. A., Davoudi, S., & Schultz, L. (2023). Thinking differently about supply chain resilience: what we can learn from social-ecological systems thinking. International Journal of Operations & Production Management, 43(1), 1-21.

Wilden, D., Hopkins, J., & Sadler, I. (2022). The prevalence of systems thinking in supply chain management: a systematic literature review. Systemic Practice and Action Research, 35(4), 491 -526.

Assessing Infrastructure Resilience in Small Islands: Methodology Proposal

Fabiane Aline Acordes
German Aerospace Center &
University of Wuppertal
fabiane.acordes@dlr.de

Sylvia Bach
Institute for Public Safety and Emergency Management, University of Wuppertal
sbach@uni-wuppertal.de

Frank Fiedrich
Institute for Public Safety and Emergency Management, University of Wuppertal
fiedrich@uni-wuppertal.de

Frank Sill Torres
Institute for Protection of Maritime Infrastructures, German Aerospace Center
frank.silltorres@dlr.de

ABSTRACT

Small islands face unique challenges when it comes to increasing the resilience of their infrastructure systems, especially against climate-related disruptions. This proposal outlines a resilience-focused methodology for assessing critical infrastructure in such environments, integrating four interrelated phases: (1) literature review and documentary analysis to understand the context and existing knowledge, (2) conducting participatory workshops for local insights, (3) semi-structured interviews for in-depth perspectives, and (4) data analysis to integrate findings. This methodology, centered on participatory approaches, seeks to explore the dynamics of the impacts of disasters on the infrastructure of small islands. Developed as part of doctoral research at the German Aerospace Center and the University of Wuppertal, this methodology will be tested on Germany's East Frisian Islands from 2023 to 2026. Its emphasis on stakeholder involvement aims to bridge gaps in existing research and provide practical strategies to enhance the long-term disaster resilience of small islands communities.

Keywords

Infrastructure resilience, small islands, participatory approach, critical infrastructure, disaster impacts.

INTRODUCTION

Recent years have witnessed significant changes in global climate patterns, resulting in increased risks of adverse weather events and environmental change (Mycoo et al., 2022). These changes have profound implications, affecting various facets of society, from local livelihoods to infrastructure (Moser et al., 2012; Roy et al., 2023).

Among the most susceptible to these impacts are small islands, which face distinct challenges due to their physical exposure and limited space and resources (Kelman, 2010; Pörtner et al., 2022; Rampengan et al., 2014). Environmental changes also pose significant threats to the stability and operability of critical infrastructure (CI), such as public water services, energy supply, communication networks and transport, which are crucial to the functioning and security of island communities.

Disruptions caused by disasters can manifest themselves in various ways, including physical damage to infrastructure assets, disruptions to essential services such as transport and utilities, and socio-economic consequences such as loss of livelihoods and population displacement. In addition, the interconnected nature of infrastructure systems exacerbates the cascading effects of disaster disruptions, amplifying their overall impact on community well-being (Labaka et al., 2014; Pal et al., 2023).

As these disasters are often unforeseen in terms of time, magnitude and location, managing the resilience of CI becomes crucially important. Within this context, resilience refers to the capacity of a system to recover from

WiP Paper – Impact-driven Emergency and Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

challenges or disruptive events (Cantelmi et al., 2021).

Managing the resilience of critical infrastructure involves considering available risk management methods to integrate different perspectives and address complexity and uncertainty (Kröger & Zio, 2011). However, small islands encounter additional obstacles when increasing infrastructure resilience due to factors such as financial constraints, geographical limitations and dependence on external aid for disaster response and recovery (Pörtner et al., 2022; Rampengan et al., 2014).

Numerous studies have explored the resilience of CIs, but only a few have focused on small islands as a distinct system. In addition, studies addressing these impacts often take a mono sectoral approach, focusing only on specific sectors and neglecting the cascading effects, dependencies, and wider social impacts of disruptions (Bruijn et al., 2019).

These issues highlight the need for innovative methodologies that not only understand the intricate dynamics of disaster disruptions to infrastructure systems but also facilitate the formulation of context-specific resilience strategies for small islands.

In addressing such challenges, participatory engagement emerges as a key function in building resilience in small island communities. By actively involving stakeholders in decision-making processes, participatory approaches benefit from local knowledge and insights, thus facilitating the development of context-specific resilience strategies (Jamero et al., 2018).

In this study, a resilience-focused methodology will be developed and proposed to analyze the impacts disaster events on CIs and societal well-being in small island contexts. By examining the impacts, duration, severity, and dependencies, the aim is to propose a methodology to measure and/or enhance resilience and safeguard the sustainability of small islands. Through a participatory approach, stakeholders are involved in developing strategies to strengthen infrastructure resilience and mitigate disaster disruptions.

CURRENT STATE OF THE ART

Disaster Impacts in the Context of Small Islands

Although there is no clear definition in the literature in terms of population or size, small islands can be broadly classified into three types (Petzold & Magnan, 2019): Small Island Developing States (SIDS), Dependent Islands, and Semi-Autonomous Subnational Island Jurisdictions (SNIJs). SIDS are recognized for their vulnerability to climate change and have been prominent in international discussions and research on this topic. Dependent islands are territories that, while not independent, are part of larger continental states and often have limited access to international aid, relying instead on national development strategies and funds. SNIJs possess a degree of autonomous government and are involved in a formal political partnership with a larger state, but they may face challenges such as limited regional cooperation and internal migration.

Despite socio-political differences, these island types share common issues related to climate change and general vulnerability due to their unique geographical and socio-economic characteristics. These impacts encompass shifts in atmospheric and oceanic conditions, leading to phenomena like sea level rise, ocean warming, ocean acidification, and alterations in extreme weather patterns such tropical cyclones and hurricanes (Nurse et al., 2014).

Extreme weather events, like hurricanes and cyclones, can inflict severe damage on small island infrastructure, including critical systems like energy, water, and sanitation (Nurse et al., 2014). Coastal infrastructure is especially vulnerable, with storm surges and high winds threatening seawalls, roads, and buildings near the shore. Limited resources for emergency response and reconstruction further exacerbate the challenges faced by these islands.

The primary effects of climate change also trigger secondary impacts, such as saltwater intrusion into groundwater and soils, marine flooding, and shifts in species distribution due to temperature and precipitation changes (Holding et al., 2016). Intrusion of saltwater can degrade water infrastructure and agricultural land, posing threats to food and water security.

From a social aspect, common challenges across small islands include limited regional-scale climate information, governance complexities, and economic vulnerabilities. These islands often rely heavily on natural resources like fisheries and tourism, making them particularly susceptible to climate-induced disruptions (Hiwasaki et al. 2014). Population dynamics, ranging from urbanization to depopulation, further strain their capacity to adapt to climate and environmental change. Moreover, small islands often depend on external financial resources and aid for development and climate change adaptation, with governance structures and historical legacies shaping their

adaptation capabilities.

Despite their vulnerabilities, small islands possess characteristics that could enhance their resilience to climate-related challenges, such as dense social networks, experience in coping with environmental adversity, and traditional knowledge (Ratter & Ratter, 2018).

CI and CI Resilience

Critical Infrastructure (CI) encompasses the essential facilities and systems crucial for the functioning of urban and rural areas, such as railways, roads, tunnels, bridges, power grids, energy sources, electricity networks, telecommunications, water supplies, sewers, etc. (Feofilovs & Romagnoli, 2017). They play a central role in society, especially when facing both technological and natural disasters.

The definition of CI is not universally standardized, as different entities and regions have varying perspectives on what constitutes CI. For instance, the U.S. Department of Homeland Security considers infrastructures critical if their incapacity would significantly impact defense and economic security (U.S Department of Homeland and Security, 2022), while the European Commission identifies CIs as those whose disruption would seriously affect the health, safety, security, or economic well-being of citizens or the functioning of governments in the European Union (EU) (EUR-Lex, 2010).

Despite the lack of a universally accepted definition, all conceptions of CI emphasize its vital role and the potentially severe societal consequences in the event of its disruption. Similarly, the term “resilience” is multidisciplinary and poses challenges in its definition. Originating in mechanics, it initially described an object's capacity to revert to its original form after undergoing external forces without fracturing (Holling, 1973). Over time, it has found broad application across various disciplines including ecology, sociology, psychology, etc. (Bruneau et al., 2003).

When applied to CIs, the concept of resilience emphasizes the ability capacity of a system to widely recover following disasters, even when infrastructure systems experience localized damage due to severe natural or human-induced disruptive occurrences (Ouyang et al., 2019). The concept is also applied to the interdependencies between CIs, recognizing that the failure of one infrastructure can have cascading effects on others (Pal et al., 2023).

CI Assessments and Participatory Approaches

Resilience is often assessed through indicators and variables that affect CIs’ organizational functionalities. They provide a means to assess baseline conditions, measure the effectiveness of interventions, and track resilience progress over time (Bruijn et al., 2019).

An indicator is factual data that can be measured, and its assessment can adopt quantitative, semi-quantitative or qualitative formats. Quantitative approaches offer standardized measures to quantify value across applications and frameworks tailored to specific domains. Semi-quantitative approaches provide numerical descriptions without complex formulas or models, while qualitative methods rely on subjective analyses and judgments, often incorporating insights from surveyed experts or operators (Cantelmi et al., 2021).

Evaluating the current resilience of a system can be accomplished through two main approaches. Firstly, by analyzing how the system responds to an event and measuring the extent of disruption or recovery across various scenarios. For example, Barabadi et al. (2020) examined the resilience of health infrastructure before and after the COVID-19 pandemic, assessing its ability to withstand and recover from the crisis.

Alternatively, resilience can be assessed by identifying the inherent characteristics of the system that promote resilience. This involves pinpointing societal factors and infrastructure qualities that contribute to resilience, such as the 100 Resilience City Framework developed by ARUP (2018). By evaluating the presence of these characteristics, analysts can gauge the system's capacity to adapt and recover from disruptive events.

Participatory processes are particularly valuable in capturing indicators in small island settings, as they facilitate the inclusion of perspectives from potentially marginalized groups into resilience planning and implementation (Bedsted and Gram, 2013). A participatory approach involves actively engaging community members in decision-making processes, ensuring that their perspectives are incorporated into the development of resilience strategies and actions (Jamero et al., 2018).

However, it's essential to recognize that labeling an approach as "participatory" does not guarantee its inclusivity at all scales, as cautioned by Cunningham (2019). This process must be accompanied by accountability and transparency to ensure that responsibility for critical infrastructure (CI) resilience is translated into practice. It

goes beyond mere consultation, emphasizing collaboration and partnership between researchers, policymakers, and local stakeholders.

This approach has been applied across various settings, as demonstrated in Talubo et al. (2022), where a combination of participatory methods was used to identify composite indicators for assessing disaster resilience in an island community.

PROPOSED METHODOLOGY

The approach proposed in this methodology involves examining the system's response to an event, quantifying both the disturbances, impacts, and dependencies, and the subsequent recovery across a specific scenario. The methodology will be employed as part of a 2023-2026 doctoral study program conducted at the University of Wuppertal and the German Aerospace Center, focusing on the East Frisian Islands in Lower Saxony-Germany as a case study. The East Frisian Islands present a pertinent context for this research due to their small island setting, vulnerability to climate change-induced hazards, and reliance on CI for socio-economic activities and community well-being.

Utilizing case studies as a research method are valuable in comprehending how individuals and communities respond to contemporary events and issues (Robson & McCartan, 2016; Yin, 2018). An exploratory approach guides the investigation of these case studies, allowing for a deeper understanding of the factors influencing priorities in the face of climate-related hazards. This approach is particularly suitable when exploring issues beyond the researcher's control, such as the unpredictable nature of climate impacts (Yin, 2018).

The study advocates a stakeholder-led approach, emphasizing the central role of participatory approaches in data collection. These methods are structured to develop in four distinct but interrelated phases, which may at times take place simultaneously:

- On Phase 1, thorough examination of existing literature, documents, reports, and academic publications an initial literature review and documentary analysis are conducted to gain a comprehensive understanding of the contextual factors and distinctive challenges encountered by small island communities, with a specific focus on the case study area.
- Phase 2 brings together stakeholders in participatory workshop(s) conducted in collaboration with local authorities, community representatives, and infrastructure managers. These will serve as interactive platforms for assessing the impacts of disaster disruptions on infrastructure systems, taking into account dependencies and cascade effects.
- On Phase 3 semi-structured interviews are conducted as a complementary approach to further enrich the insights gained from other phases, particularly regarding stakeholders' perceptions and experiences.
- Phase 4 employs an integrated qualitative and/or quantitative analysis based on the findings from the scenario-based impact assessments in order to refine strategies for mitigating disruptions and enhancing CI resilience.

Phase One: Literature Review and Documentary Analysis

The methodology employed in this study encompasses an extensive review of existing literature on disaster impacts, infrastructure resilience assessments, and participatory approaches in small island contexts. This review aims to identify gaps and relevant concepts pertinent to the subject matter, focusing on the specific context and requirements of the small island(s) to be assessed.

To conduct the literature review, a systematic approach is adopted, involving keyword searches, content analysis, and the examination of titles and abstracts to identify relevant literature. This approach ensures a comprehensive understanding of the current state of knowledge in the field, allowing for the identification of key themes, trends, and areas requiring further investigation. The review encompasses a range of sources, including scientific studies, case reports, and institutional documents, focusing on climate change vulnerability and resilience in small island contexts.

In addition to the literature review, a documentary analysis method is employed for the study case(s). This qualitative technique involves the analysis of content within various documents, such as official reports, meeting minutes, newspaper articles, social media posts, and other written or recorded content. Documentary analysis is valuable for gaining deeper insights into the research topic, complementing research efforts (Robson & McCartan, 2016).

The records produced during documentary analysis provide a deeper understanding of the case studies as real-world communities, serving as a vital step before fieldwork begins. Specific key questions addressed include which scenarios, key stakeholders, and elements of infrastructure should be included in the assessment process. An example of potential CI sectors to be considered is presented in Table 1 (Petrakos & Kotzanikolaou, 2019).

Table 1: Example of CI sectors and their respective service/organizations.

CI Sector	Service/Organizations
Transport	- Air/road/train/maritime transport
	- Services and network
Water (supply and disposal)	- Storage, distribution, quality
	- Collection and treatment
Energy (electricity, petroleum, gas, etc.)	- Production, transport and distribution
Health	- Emergency healthcare
	- Hospital care
	- Medical supplies
	- Nursing care
Food	- Agriculture/fishery production
	- Food supply and distribution
Information technologies & media	- Web/cloud services
	- Internet, voice/data communication
	- Media press and broadcasting
Administration	- Government functions (economy, engineering, etc.)
Defense and civil protection	- Crisis management, emergency and rescue services
Tourism	- Touristic infrastructure (property, visitors' sites, recreational services)
	- Sector employment
Housing and land use	- Village urban planning zone
	- Private and public properties
Community and social organizations	

Phase Two: Participatory Workshops

Participatory workshops are conducted on the selected small islands to engage communities in assessing potential impacts and cascading effects on infrastructure using developed scenarios.

Scenarios serve as a crucial tool to address uncertainty, particularly concerning poorly understood, difficult-to-quantify, or unidentified risks. They provide a systematic method for assessing how diverse and complex risks could impact society, essentially evaluating the resilience of systems to potential disruptions (Strong et al., 2020).

During the workshops, collaboration between various stakeholder groups, including community members, municipal representatives, critical infrastructure operators, authorities and organizations with security tasks, is encouraged to address concrete indicators of functionality in crisis events and dependencies on other organizations. This participatory approach promotes dialogue, knowledge exchange, and collective problem-solving (Bizikova et al., 2011; Sheppard et al., 2011), and has also been recorded in the particular context of coastal communities (Simpson et al., 2012). Peer learning is also promoted, as workshops lead to a facilitated exchange of best practices and experiences between participants.

The workshops are structured to commence with an explanation of the workshop's purpose and an introductory keynote addressing key issues faced by small islands during disasters, emphasizing the significance of evaluating their impacts on infrastructure and population.

Then, participants are presented with a range of potential scenarios, covering sudden-onset hazards ranging from

natural occurrences such as floods, coastal erosion, droughts or storms, to human-induced events such as blackouts, disease epidemics or civil conflicts (Strong et al., 2020). These scenarios are discussed among the participants and ranked based on their potential impacts and occurrence probability in their small island context.

In the second part, one scenario is selected for further analysis. Stakeholders identify emerging issues from the selected scenario for their respective facilities and interests, explore mutual dependencies, and support potentials between different organizations. Questions to stakeholders revolve around indicators: variables that affect their organization's functionalities (e.g. number of accommodation facilities or access sea routes into the island at a given time). Questions are exemplified below:

- What indicators are necessary in your organization to ensure its functionality? Give examples in measurable and comprehensive units.
- How can those indicators be impacted in each disaster scenario?
- What dependencies exist with other organizations? What possible impact does the non-conformity of these dependencies have on the infrastructure system and population?
- What challenges do you foresee arising in your organization from each scenario?

The results are collectively discussed, and the points raised by the participants can be listed and placed in diagrams for better visualization of the CIs and their dependencies. After the discussion, a feedback session takes place to gather additional insights and perspectives.

Phase Three: Semi-Structured Interviews

Building upon workshop insights, interviews are conducted with identified actors from the study area(s) to gather valuable insights into various aspects of the island's infrastructure and also to complement the data and results of the other phases.

The semi-structured interview provides consistency in conversations through predefined prompts, while also allowing for unexpected topics of interest to be explored and for clarifications/follow-up questions from participants' answers (Paine, 2015).

These interviews are conducted with experts who possess relevant knowledge or access to such knowledge within their organizational units. The insights provided by these experts offer valuable perspectives on different facets of the island's infrastructure, contributing to a comprehensive understanding of its strengths, weaknesses, and resilience.

Considering the scenario chosen during the workshops, a second part of the interview focuses on what kind of mitigating actions could be implemented by the organization or sector itself or in collaboration with others.

Interview questions and dynamics are flexibly structured to ensure alignment with the research objectives and methodology. This flexibility allows for the exploration of diverse topics related to infrastructure resilience, climate change impacts, and disaster preparedness. It also facilitates the first hand transmission of participant knowledge and expertise, enabling researchers to glean nuanced insights into the challenges and opportunities facing the island's infrastructure.

Phase Four: Data Analysis

Data obtained from phases 1, 2 and 3 can be examined using quantitative or qualitative approaches. The debate over which method to use when studying complex issues is tenuous, as each approach has its own strengths and weaknesses (Matsika et al., 2016).

Decision makers often favor quantitative data for their decision-making processes, especially in analyses like cost-benefit assessments that require numerical values (Matsika et al., 2016). However, qualitative assessments offer valuable context and understanding, complementing quantitative data (Bernroider et al., 2016).

For instance, Van Laere et al. (2017) investigated challenges encountered during disruptions in the payment system using inductive qualitative research methods. Through document analysis, interviews with key sector representatives, and workshops with local and national stakeholders, they developed scenarios and identified resilience challenges for CIs. Their analysis underscored the importance of addressing communicative challenges such as maintaining trust and preventing panic among the public, alongside technical solutions.

However, Fekete & Fiedrich (2018) expressed skepticism about measuring resilience quantitatively due to the complexity of CIs but also cautioned against relying solely on qualitative assessments. They advocated for a

mixed-methods approach, combining quantitative and qualitative dimensions to provide a broader perspective and cater to diverse end users. Münzberg et al. (2017), exemplifies this approach on supporting community resilience through a spatial-temporal vulnerability assessment. By integrating various indicators using the Delphi method and Monte Carlo simulation, they provided decision-makers with a comprehensive understanding of power outage impacts, aiding in resilience-building efforts.

In the context of disaster resilience and CI, the data analysis techniques may involve utilizing quantitative methods to analyze numerical data collected during impact assessments or participatory workshops. This could include statistical analysis to identify patterns or trends in the data.

Qualitative data analysis techniques, on the other hand, may involve thematic analysis of interview transcripts or documentary analysis findings. This qualitative approach allows researchers to explore the nuanced perspectives and experiences of stakeholders, providing rich insights into their perceptions and behaviors.

CONCLUSION

The resilience-focused methodology proposed in this study offers a comprehensive structure for assessing infrastructure resilience in small island environments. By synthesizing findings from literature reviews, documentary analyses, interviews and participatory workshops the methodology aims to capture the complex dynamics of disaster impacts and cascading effects within these vulnerable contexts.

One key feature of the proposed methodology is its emphasis on scenario-based participatory approaches, which enable stakeholders to actively engage in the assessment process and explore various resilience strategies. Through this collaborative approach, the methodology seeks to provide a nuanced understanding of how disasters disrupt infrastructure systems and the subsequent consequences on communities.

Moreover, the discussion surrounding the choice between quantitative and qualitative approaches to analyze the data highlights the flexibility inherent in the proposed methodology. By acknowledging the advantages and drawbacks of each approach, the methodology offers researchers and organizations a practical guide for tailoring their assessment strategies to the unique characteristics of small island settings. Overall, the application of this methodology may generate contributing results for advancing knowledge on infrastructure resilience and supporting efforts for mitigating the impacts of disasters and safeguarding the sustainability of small islands globally.

REFERENCES

Arup Group Limited (2018) City resilience index: understanding and measuring city resilience.

Barabadi, A., Ghiasi, M. H., Nouri Qarahasanlou, A., & Mottahedi, A. A holistic view of health infrastructure resilience before and after COVID-19. ABJS, vol. 8, no. Covid-19 Special Issue (2020).

Bedsted, B., & Gram, S. (2013). Participatory climate change adaptation in Kalundborg, Denmark. *Climate Change Adaptation in Practice: From strategy development to implementation*, 11-23.

Bernroider, E. W., Margiol, S., & Taudes, A. (2016). Towards a General Information Security Management Assessment Framework to Compare Cyber-Security of Critical Infrastructure Organizations. In *Research and Practical Issues of Enterprise Information Systems: 10th IFIP WG 8.9 Working Conference, CONFENIS 2016, Vienna, Austria, December 13–14, 2016, Proceedings 10* (pp. 127-141). Springer International Publishing.

Bizikova, L., Burch, S., Robinson, J., Shaw, A., & Sheppard, S. (2011). Utilizing Participatory Scenario-Based Approaches to Design Proactive Responses to Climate Change in the Face of Uncertainties. In G. Gramelsberger & J. Feichter (Eds.), *Climate Change and Policy: The Calculability of Climate Change and the Challenge of Uncertainty* (pp. 171–190). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-17700-2_8.

Bruijn, K. M. de, Maran, C., Zygnerski, M., Jurado, J., Burzel, A., Jeuken, C., & Obeysekera, J. (2019). Flood Resilience of Critical Infrastructure: Approach and Method Applied to Fort Lauderdale, Florida. *Water*, *11*(3), 517. https://doi.org/10.3390/w110 30517.

Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., Shinozuka, M., Tierney, K., Wallace, W. A., & Winterfeldt, D. von (2003). A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthquake Spectra*, *19*(4), 733–752. https://doi.org/10.1193/1.1623497.

Cantelmi, R., Di Gravio, G., & Patriarca, R. (2021). Reviewing qualitative research approaches in the context of critical infrastructure resilience. *Environment System and Decisions*, *41*(3), 341–376. https://doi.org/10.1007/s10669-020-09795-8.

Cunningham, F. (2019). *Adaptation to the impacts of climate change in small island communities: an analysis of Scottish case studies* (Doctoral dissertation, University of St Andrews).

EUR-Lex, Critical infrastructure protection, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=legissum %3Al33259, 2010 (accessed 05 Jun. 2023).

Fekete, A., & Fiedrich, F. (Eds.). (2018). *The Urban Book Series*. *Urban Disaster Resilience and Security: Addressing Risks in Societies* (1st ed. 2018). Springer International Publishing; Imprint: Springer.

Feofilovs, M., & Romagnoli, F. (2017). Resilience of critical infrastructures: probabilistic case study of a district heating pipeline network in municipality of Latvia. *1876-6102*, *128*, 17–23. https://doi.org/10.1016/j.egypro.2017.09.007.

Hiwasaki, L., Luna, E., Syamsidik, & Marçal, J. A. (2015). Local and indigenous knowledge on climate-related hazards of coastal and small island communities in Southeast Asia. *Climatic Change*, *128*, 35-56.

Holding, S., Allen, D. M., Foster, S., Hsieh, A., Larocque, I., Klassen, J., & Van Pelt, S. C. (2016). Groundwater vulnerability on small islands. *Nature Climate Change*, *6*(12), 1100-1103.

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual review of ecology and systematics*, *4*(1), 1-23.

Jamero, M. L., Onuki, M., Esteban, M., & Tan, N. (2018). Community-based adaptation in low-lying islands in the Philippines: Challenges and lessons learned. *Regional Environmental Change*, *18*(8), 2249 – 2260. https://doi.org/10.1007/s10113-018-1332-8.

Kelman, I. (2010). Hearing local voices from Small Island Developing States for climate change. *Local Environment*, *15*(7), 605–619. https://doi.org/10.1080/1354 9839.2010.498812.

Kröger, W., & Zio, E. (2011). Conceptual Frameworks for Vulnerability Assessment. In W. Kröger, E. Zio, & M. Schläpfer (Eds.), *Vulnerable systems* (pp. 55–64). Springer. https://doi.org/10.1007/978-0-85729-655-9_5.

Labaka, L., Comes, T., Hernantes, J., Sarriegi, J. M., & Gonzalez, J. J. (2014). Implementation Methodology of the Resilience Framework. In *2014 47th Hawaii International Conference on System Sciences*. IEEE. https://doi.org/10.1109/hicss.2014.26.

Matsika, E., O'Neill, C., Battista, U., Khosravi, M., Laporte, A. d. S., & Munoz, E. (2016). Development of Risk Assessment Specifications for Analysing Terrorist Attacks Vulnerability on Metro and Light Rail Systems.

2352-1465, *14*, 1345–1354. https://doi.org/10.1016/j.trpro.2016.05.207.

Moser, S. C., Jeffress Williams, S., & Boesch, D. F. (2012). Wicked Challenges at Land's End: Managing Coastal Vulnerability Under Climate Change. *Annual Review of Environment and Resources*, *37*(1), 51–78. https://doi.org/10.1146/annurev-environ-021611-135158.

Münzberg, T., Wiens, M., & Schultmann, F. (2017). A spatial-temporal vulnerability assessment to support the building of community resilience against power outage impacts. *Technological Forecasting and Social Change*, *121*, 99–118. https://doi.org/10.1016/j.techfore.2016.11.027.

Mycoo, M., Wairiu, M., Campbell, D., Duvat, V., Golbuu, Y., Maharaj, S., Nalau, J., Nunn, P., Pinnegar, J., & Warrick, O. (2022). Small islands. *IPCC WGII Sixth Assessment Report.*

Nurse, L. A., McLean, R. F., Agard, J., Briguglio, L. P., Duvat-Magnan, V., Pelesikoti, N., ... & Webb, A. (2014). Small islands. *Climate change 2014: Impacts, adaptation, and vulnerability. Part B: Regional aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change*, pp-1613-1654. https://doi.org/10.2134/jeq2008.0015br.

Ouyang, M., Liu, C., & Xu, M. (2019). Value of resilience-based solutions on critical infrastructure protection: Comparing with robustness-based solutions. *Reliability engineering & system safety*, *190*, 106506.

Paine, G. (2015). A Pattern-Generating Tool for Use in Semi-Structured Interviews. *The Qualitative Report.* Advance online publication. https://doi.org/10.46743/ 2160-3715/2015.2123.

Pal, I., Kumar, A., & Mukhopadhyay, A. (2023). Risks to Coastal Critical Infrastructure from Climate Change. *Annual Review of Environment and Resources*, *48*(1), 681–712. https://doi.org/10.1146/annurev-environ-112320-101903.

Petrakos, N., & Kotzanikolaou, P. (2019). Methodologies and Strategies for Critical Infrastructure Protection. *Critical Infrastructure Security and Resilience*, 17–33. https://doi.org/10.1007/978-3-030-00024-0_2.

Ratter, B. M., & Ratter, B. M. (2018). Island vulnerability and resilience. *Geography of small islands: Outposts of globalisation*, 173-199. https://doi.org/10.1007/978-3-319-63869-0_6.

Petzold, J., & Magnan, A. K. (2019). Climate change: thinking small islands beyond Small Island Developing States (SIDS). *Climatic change*, 152(1), 145-165.

Pörtner, H. O., Roberts, D., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama-Phd, B., Belling, D., Dieck, W., Götze, S., Kersher, T., Mangele, P., Maus, B., Mühle, A., & Weyer, N. (2022). *Climate Change 2022: Impacts, Adaptation and Vulnerability Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* https://doi.org/10.1017/9781009325844.

Rampengan, M. M. F., Boedhihartono, A. K., Law, L., Gaillard, J. C., & Sayer, J. (2014). Capacities in Facing Natural Hazards: A Small Island Perspective. *International Journal of Disaster Risk Science*, *5*(4), 247–264. https://doi.org/10.1007/s13753-014-0031-4.

Robson, C., & McCartan, K. (2016). Real World Research: A Resource for Users of Social Research Methods in Applied Settings. 4th ed. John Wiley and Sons Ltd.; West Sussex, UK: 2016.

Roy, P., Pal, S. C., Chakrabortty, R., Chowdhuri, I., Saha, A., & Shit, M. (2023). Effects of climate change and sea-level rise on coastal habitat: Vulnerability assessment, adaptation strategies and policy recommendations. *Journal of Environmental Management*, *330*, 117187. https://doi.org/10.1016/j.jenvman.2022.117187.

Sheppard, S. R., Shaw, A., Flanders, D., Burch, S., Wiek, A., Carmichael, J., Robinson, J., & Cohen, S. (2011). Future visioning of local climate change: A framework for community engagement and planning with scenarios and visualisation. *0016-3287*, *43*(4), 400–412. https://doi.org/10.1016/j.futures.2011.01.009.

Simpson, M., Clarke, J., Scott, D., New, M., Karmalkar, A., Day, O., Taylor, M., Gossling, S., Wilson, M., Chadee, D., Stager, H., Waithe, R., Stewart, A., Georges, J., Hutchinson, N., Fields, N., Sim, R., Rutty, M., Matthews, L., & Agosta G'meiner, A. (2012). *THE CARIBSAVE CLIMATE CHANGE RISK ATLAS (CCCRA) - Climate Change Risk Profile for the Dominican Republic.*

Strong, K., Carpenter, O., & Ralph, D. (2020). Scenario best practices: developing scenarios for disaster risk reduction. *Cambridge Centre for Risk Studies at the University of Cambridge*, 12–15.

Talubo, J. (2022). *Assessing small island socio-ecological resilience for disaster risk and recovery planning in the Philippines* (Doctoral dissertation, University of Surrey).

U.S Department of Homeland and Security, 2022. Critical Infrastructure Security & Resilience Research (CISRR), Fact Sheet https://www.dhs.gov/sites/default/files/2023-06/23_0614_st_Critical_Infrastructure _Security_Resilience_Research.pdf (accessed 12 Aug. 2023).

Van Laere, J., Berggren, P., Gustavsson, P., Ibrahim, O., Johansson, B., Larsson, A., Lindqwister, T., Olsson, L., & Wiberg, C. (2017). Challenges for critical infrastructure resilience: Cascading effects of payment system disruptions. In (pp. 281–292). ISCRAM. https://www.diva-portal.org/smash/record.jsf?pid= diva2:1094917.

Yin, R. K. (2018). Case study research and applications (Vol. 6). Thousand Oaks, CA: Sage.

A Study on The Heterogeneity of Tropical Cyclone Vulnerability Functions in Different Provinces along The Coastal Areas in China

Chenning Pan
Joint International Research Laboratory of Catastrophe Simulation and Systemic Risk Governance, Beijing Normal University, Zhuhai 519087, China;
School of National Safety and Emergency Management, Beijing Normal University, Zhuhai 519087, China
202321099015@mail.bnu.edu.cn

Xiaoyong Ni*
Joint International Research Laboratory of Catastrophe Simulation and Systemic Risk Governance, Beijing Normal University, Zhuhai 519087, China;
School of National Safety and Emergency Management, Beijing Normal University, Zhuhai 519087, China
nixiaoyong@bnu.edu.cn

Xifan Shen
Bay Area International Business School, Beijing Normal University, Zhuhai 519087, China
202111059036@mail.bnu.edu.cn

Ziyi Peng
Bay Area International Business School, Beijing Normal University, Zhuhai 519087, China
2960830241@qq.com

*Correspondence: nixiaoyong@bnu.edu.cn

ABSTRACT

Coastal regions in China have long been plagued by tropical cyclone (TC) disasters, and TC vulnerability function for disaster-bearing bodies is a widely used quantitative vulnerability assessment method in face of TCs. TC vulnerability functions may vary across different regions due to geographical heterogeneity, hindering this method from universal use. This study systematically collects a 15-year-long dataset from 12 provinces in China. Six alternative TC vulnerability functions are comprehensively investigated by province, and a library of the optimal TC vulnerability functions for three kinds of disaster-bearing bodies, that is, economic, population, and crop in different provinces is developed. The heterogeneity of TC vulnerability in different provinces along the coastal areas in China is then thoroughly analyzed. This paper provides a deeper understanding of the spatial heterogeneity characteristics of disasters and disaster-bearing bodies, and provides good references for local departments to carry out more targeted disaster prevention and reduction work.

Keywords

Tropical cyclone, vulnerability functions, coastal areas in China, disaster loss assessment.

INTRODUCTION

Tropical cyclones (TCs) are low-pressure vortices that occur over tropical or subtropical oceans. A TC with a central sustained wind speed of 32.7 to 41.4 meters per second is named as a typhoon (located in the western North Pacific) or a hurricane (located in other typhoon-prone regions) (WMO, 2022; National Meteorological

WiP Paper – Impact-driven Emergency and Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Information Center, 2023). It is reported that TCs are natural disasters that have a wide range of impacts and are extremely destructive. Over the past twenty years, there have been a total of about two thousand tropical storms globally with a wide distribution in Asia, North America and Europe. The number of people affected by TCs accounted for 19.52%; the number of deaths and missing persons accounted for 19.17%; and the direct economic losses accounted for 48.57% of the total global natural disasters (Academy of Disaster Reduction and Emergency Management et al., 2023). Relevant data show that the TC disaster is one kind of the main natural disasters faced by human beings, which needs great and urgent attention.

The area of China's coastal provinces and provinces adjacent to coastal provinces is a typical region most severely hit by typhoons (Xu et al., 2015; Wan et al., 2023). Every year, a large number of typhoons make landfall in the southeastern coastal areas of China, causing huge losses. In August 14th, 2019, typhoon "Likima" affected a total of 14.02 million people in China with 2.10 million people urgently relocated, and caused 57 deaths, as well as a direct economic loss of 53.72 billion yuan. In September 5th, 2023, typhoon "Haikui" had caused record-breaking extreme rain in Fujian, Guangdong, and other places. Many findings also suggest that typhoon disasters in coastal areas in China will become increasingly severe in the future (Lee et al., 2020; Wang et al., 2021). Therefore, it is of great significance to study typhoon disasters in China's coastal areas.

In this study, the classic method of typhoon vulnerability functions is utilized to depict the current typhoon risk in China's coastal area. It is widely recognized that typhoon vulnerability functions have significant spatial heterogeneity. This paper currently focuses on the heterogeneity of typhoon vulnerability functions in different provinces along the coastal areas in China. Considering that the loss of typhoon disasters not only depends on the typhoon intensity itself, but also depends on the resilience characteristics of disaster-bearing bodies and the scientific nature of disaster prevention measures, the authors believe that the results provided by this paper have an important reference value for improving emergency management capabilities and strategies in future emergency work in face of typhoons.

The remainder of the paper is organized as follows. Section 2 is the literature review, which reviews the application areas of vulnerability curve methods and their specific applications in typhoon loss assessment. Section 3 outlines the writing approach of the article. Section 4 introduces the research area and data sources used in the paper. Section 5 presents the form of vulnerability curves selected for this study. Section 6 analyzes the heterogeneity of vulnerability curves in the study area. Section 7 discusses the limitations of the research and future directions. Finally, Section 8 summarizes the main findings and their significance.

LITERATURE REVIEW

Vulnerability and Vulnerability Curve

With the concept of "vulnerability" being widely used in multiple fields during the past decades, such as politics (White at al., 2018), economic (Barnett, 2020), ecosystem (Xu et al., 2023), infectious disease (Shifa et al., 2022), and disasters (Colombi et al., 2008; Godfrey et al., 2015), etc., more and more scholars in the field of disaster studies began to pay attention to the technology of vulnerability functions, which could be rapidly assessed the risk. The vulnerability function, also known as the vulnerability curve, is a quantitative assessment method depicting the relationship between the intensity of hazard factors and the loss degree of the disaster-bearing bodies. The vulnerability function (curve) is usually represented in a two-dimensional coordinate system, where the horizontal axis represents the hazard intensity, such as rainfall intensity or wind speed, etc., and the vertical axis represents the loss degree of the disaster-bearing bodies, such as the loss rate of crops, buildings, or economic, etc. (Zhang et al., 2016).

The Application of Vulnerability Curves/Functions in the Field of Disasters

The vulnerability curve has been widely applied across various fields in recent years. It has become a crucial component in disaster assessment, quantitative risk analysis, and the creation of risk maps. Zhou et al. summarized the research progress of vulnerability curves for various disasters, including floods, earthquakes, landslides, droughts, typhoons (hurricanes) (Zhou et al., 2012). Monteleone et al. reviewed 52 articles on vulnerability curves of different crops under extreme weather events and climate change, outlining the vulnerability curve forms of different crops under conditions such as droughts, floods, and cold weather (Monteleone et al., 2023). Papathoma-Köhle et al. summarized the three analysis methods and advantages and disadvantages of physical vulnerability analysis for buildings based on debris flow disaster, including vulnerability matrices, curves, and indicators (Papathoma-Köhle et al., 2017). Vamvatsikos et al. reviewed existing concepts and methods for assessing structural vulnerability to natural disasters (earthquakes, landslides, tsunamis, and wind disasters) from a structural

and engineering perspective (Vamvatsikos et al., 2010). This indicates that the application scope of vulnerability curves in the disaster domain is wide, with diverse forms and a rich variety of vulnerability assessment targets covered.

The Current Status of Vulnerability Curves in TC Risk Assessment and Loss Analysis

Nowadays, TC vulnerability function/curve method has been a widely recognized and effective method in TC risk assessment and loss analysis. There are usually two methods to establish the typhoon vulnerability function/curve - one is the engineering experiment (simulation) method, and the other is the field investigation (data inversion) method. As to the engineering experiment (simulation) method, the techniques like wind tunnel test and computer simulation are usually used to detect under what TC intensity a whole building or structural parts will fail; then, based on the original value or the reconstruction cost of the structure, the loss rate of the structure can be calculated (Zhong et al., 2017). For example, the Hazus model and the Florida hurricane model applied by the U.S. government used the structural engineering simulation to build the building vulnerability (Pinelli et al., 2011; Vickery et al., 2006a, 2006b). Compared with the engineering experiment (simulation) method, the field investigation (data inversion) method, which is based on real historical data record, is usually able to obtain data closer to the real result. Acosta et al. assessed typhoon damage in Philippine schools using empirical vulnerability curves, based on damage quantification against maximum local wind speeds (Acosta et al., 2018). Zhu et al. developed a road vulnerability model against TCs in Hainan Province, using historical road damage data and historical TC intensity measures like precipitation and wind speed (Zhu et al., 2022).

Despite numerous studies on TC vulnerability functions, it should be noted that TC vulnerability functions often differ across regions due to the geographical spatiotemporal heterogeneity. Limited by the availability of research data, many studies only focus on a specific province or city without comparing differences between different regions. In some practical applications, there have been cases where the typhoon vulnerability curve of a certain province has been used to assess losses in another province when the historical data records of the latter is inaccessible, introducing significant human error into typhoon damage assessments. Therefore, developing a library of typhoon vulnerability functions suitable for different regions has become a worthwhile hotspot in this field.

THE OVERALL STRUCTURE

The overall structure of this study is shown in Figure 1. The region studied comprises twelve coastal provinces and provinces adjacent to coastal provinces in China, having a characteristic of being vulnerable to typhoons. The basic data by province and a 15-year-long typhoon dataset in all provinces, including disaster intensity and losses caused by TCs, are collected. Based on the basic dataset, six alternative vulnerability functions are calibrated to fit the loss rates of the economic, population, and farmland in each province, respectively, and finally, the optimal vulnerability functions are selected to profile the losses caused by typhoons. The heterogeneity of the losses of different disaster-bearing bodies in different provinces along the coastal areas in China are analyzed. The results provided by this paper could provide a profound understanding of typhoon disasters in coastal areas in China, and also the reference and technological support for the government emergency management to cope with typhoon disasters.

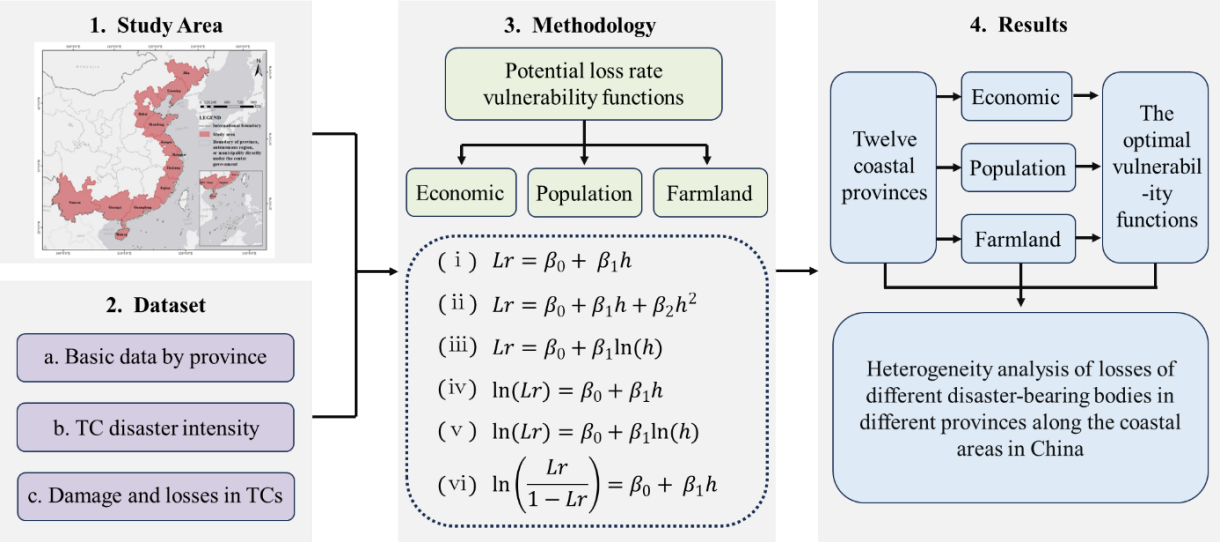


Figure 1. The Overall Structure of This Study

STUDY AREA AND DATASETS

Study Area

"China Marine Statistical Yearbook" defines coastal areas as regions with coastlines, including both categories of mainland coastlines and island coastlines (Ministry of Natural Resources of the People's Republic of China, 2017). According to this definition, ten coastal provinces in China are considered in this paper, including Liaoning, Hebei, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, Guangxi, and Hainan. According to the historical records of typhoon-related disasters and the impact of typhoons, Jilin Province and Yunnan Province are also tagged as provinces prone to typhoons, so these two provinces are also included in the scope of this study. Due to the data availability issue, Tianjin Province, Taiwan Province, Hong Kong Special Administrative Region (SAR) and Macao SAR are temporarily not taken into account in this paper. Summarily, this study includes 12 provinces in total. Figure 2 illustrates the study area in this study.

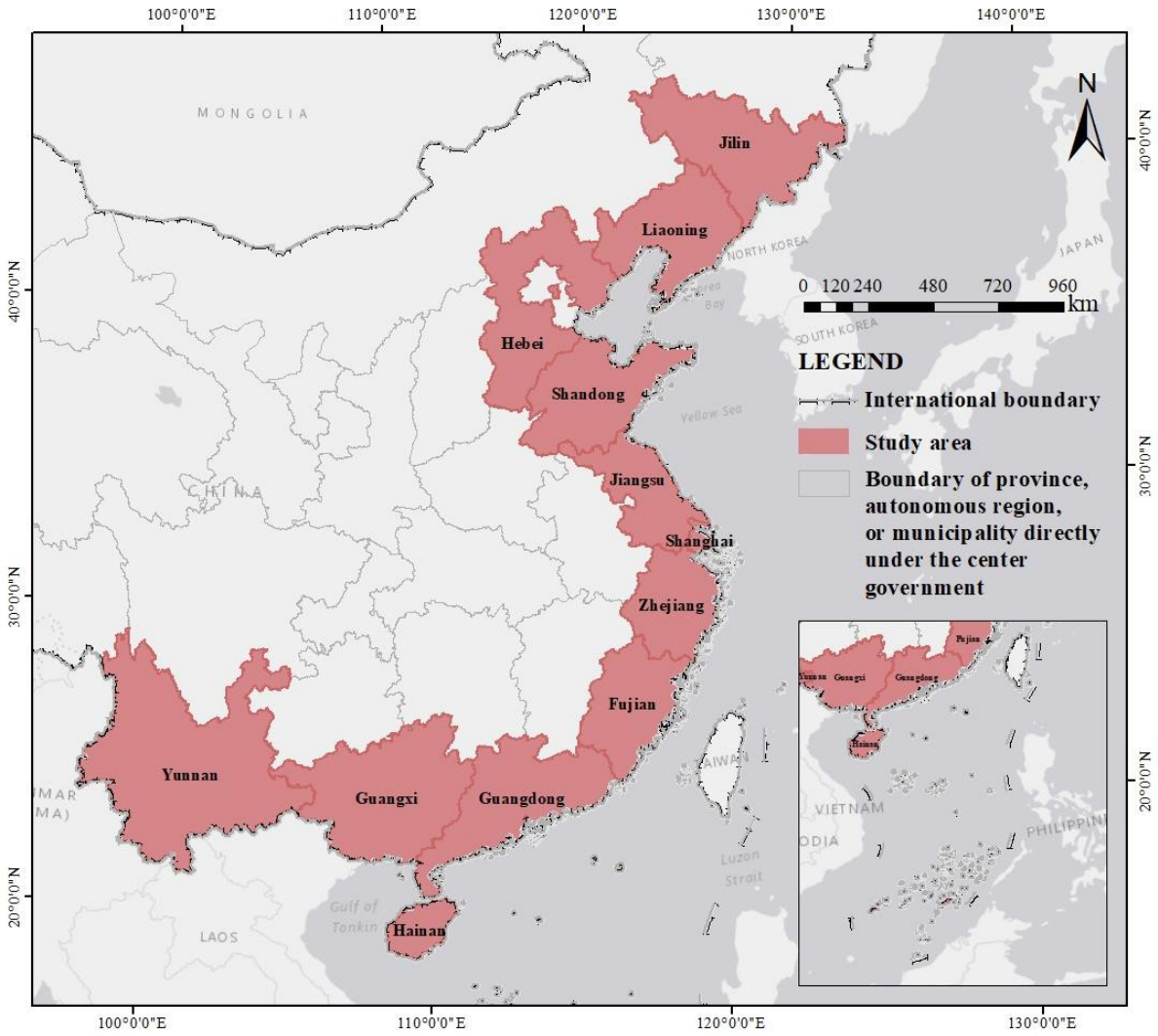


Figure 2. The Study Area in This Study

The basic situations of all the provinces involved in this study are shown in Table 1. It should be noted that the number of meteorological stations in Table 1 refers to the number of the stations which have ever reported the total cumulative rainfall during the typhoons. Stations that do not give records are not counted. The precipitation data from multiple meteorological stations consider the spatial heterogeneity of total rainfall in each province. A total of 451 typhoons with recorded disasters occurred in 22 provinces in China from 2003 to 2018. Among them, the 12 provinces selected for this study experienced a total of 357 typhoon events, indicating that this region is under serious threat of typhoon disasters.

Table 1. Basic Overview of The Study Area (Using Data in 2018)

Province	Gross Regional Product (100 million yuan)	Population (10000 persons)	Total Sown Area (1000 hectares)	Number of meteorological stations	Number of typhoons from 2003 to 2018
Jilin	15074.6	2704	6080.9	6	5
Liaoning	25315.4	4359	4207.1	5	12
Hebei	36010.3	7556	8197.1	5	9
Shandong	76469.7	10047	11076.8	6	16
Jiangsu	92595.4	8051	7520.2	7	21
Shanghai	32679.9	2424	282.3	1	17
Zhejiang	56197.2	5737	1978.7	5	44
Fujian	35804	3941	1577.3	3	59

Guangdong	97277.8	11346	4279.4	7	69
Guangxi	20352.5	4926	5972.4	6	45
Hainan	4832.05	934	712.9	2	45
Yunnan	17881.1	4830	6890.8	8	15

Datasets

Since the precipitation data for each typhoon is only available until 2018, this study covers a period of 15 years from 2003 to 2018. The basic data by province, such as gross regional product (GRP), population, and sown area are obtained from the "China Statistical Yearbook" (National Bureau of Statistics of China, 2004-2019). The total precipitation caused by individual typhoons, a factor contributing to disasters, is extracted from the Tropical Cyclone Data Center of the China Meteorological Administration (Ying et al., 2014; Lu et al., 2021). The typhoon disaster loss data, including direct economic losses, affected population, and affected crop area, are sourced from the "Yearbook of Meteorological Disasters in China" (China Meteorological Administration, 2004-2019). All the involved data is publicly available, and the authors have placed download links of all the data in References section.

METHODOLOGY

The alternative forms of the TC vulnerability functions considered in this paper are inspired by the work by Jiang et al. (Jiang et al., 2023). The general form of the vulnerability regression model with the loss rate as the dependent variable is as follows:

$$g(Lr) = f(h) + \varepsilon, \varepsilon \sim N(0, \sigma^2) \quad (1)$$

Where Lr denotes the loss rate; $g(Lr)$ denotes the transformation function of the dependent variable Lr ; h denotes the hazard intensity that serves as the independent variable; $f(h)$ denotes the regression function of the independent variable h. The function $f(h)$ may depend on several parameters, denoted by β_0, β_1, etc., which remain to be estimated. ε denotes the random error term, usually conforming to the normal distribution.

Vulnerability regression models can be categorized into three main types: linear models, logarithmic models and probability curve model.

Linear Models

The linear model does not transform the dependent variable (the loss rate) to other mathematical forms. That is, $g(Lr) = Lr$. The linear models comprise three forms, namely simple linear model, polynomial model, linear-logarithmic model. The formulas are as follows, respectively.

$$Lr = \beta_0 + \beta_1 h \quad (2)$$

$$Lr = \beta_0 + \beta_1 h + \beta_2 h^2 \quad (3)$$

$$Lr = \beta_0 + \beta_1 \ln(h) \quad (4)$$

where: β_1 denotes the slope parameter, indicating the degree and trend of the change in the loss rate with change in hazard intensity. β_0 denotes the intercept parameter, representing the baseline loss rate when the hazard indicator is zero. In formula (3) β_2 denotes quadratic term. It should be noted that, although a non-zero intercept term is challenging to interpret, β_0 is not forced to be zero in this study. The main reason for this is that the cumulative rainfall intensity during a typhoon, which is selected as an indicator for hazard intensity, may not paint the full picture of that typhoon. Therefore, setting the parameter β_0 artificially to zero may not make sense.

Logarithmic Models

Logarithmic models involve the operation of taking the logarithm of the dependent variable Lr, that is, $g(Lr) = \ln(Lr)$. The logarithmic models comprise two forms, namely logarithmic-linear model and logarithmic-logarithmic model. The formulas are as follows, respectively.

$$\ln(Lr) = \beta_0 + \beta_1 h \quad (5)$$

$$\ln(Lr) = \beta_0 + \beta_1 \ln(h) \quad (6)$$

where: the slope parameter β_1 in Formula (5) and Formula (6) represents the absolute change and the relative

change in hazard intensity, respectively.

Probability Curve Model

Linear models and logarithmic models cannot guarantee that the predicted values of the loss rate will fall entirely within the interval of [0, 1]. In order to overcome this shortcoming, another function form is introduced here which takes the logarithm of the ratio of the loss rate Lr to $(1 - Lr)$. This model is named as logistic curve model. The formula is as follows.

$$\ln\left(\frac{Lr}{1 - Lr}\right) = \beta_0 + \beta_1 h \tag{7}$$

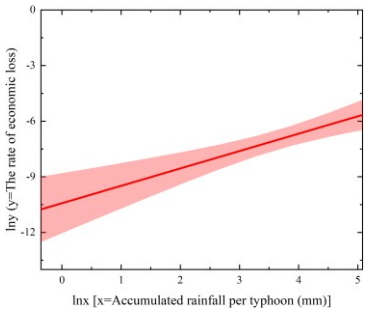
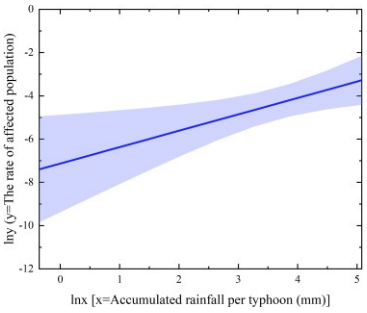
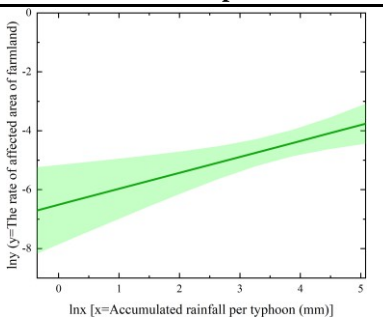
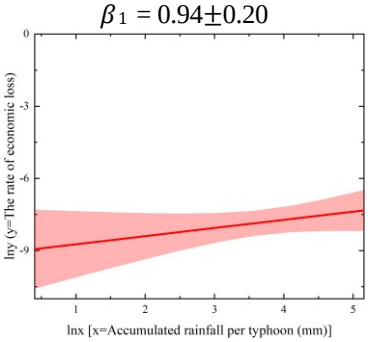
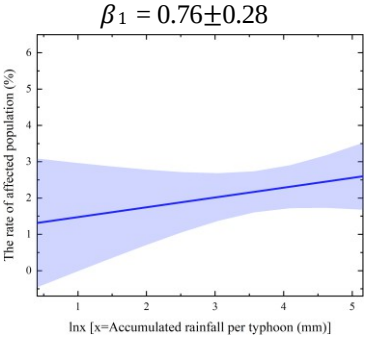
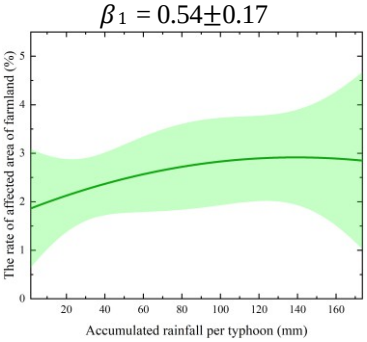
RESULTS

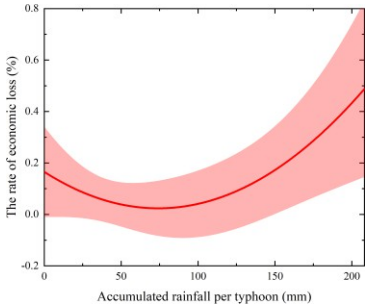
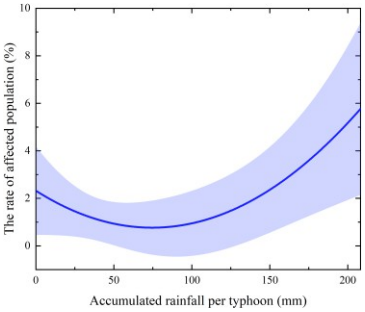
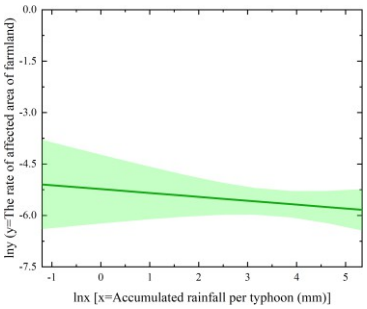
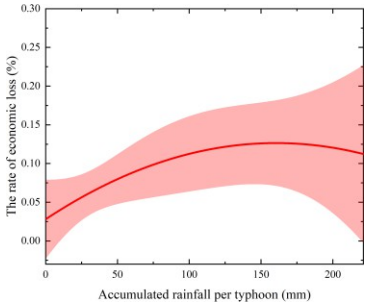
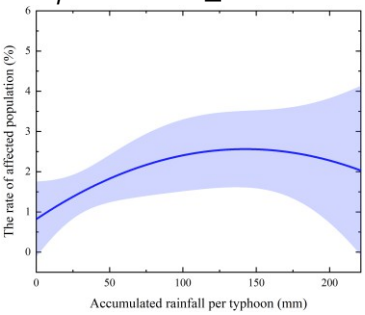
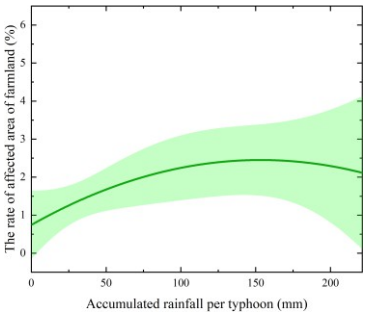
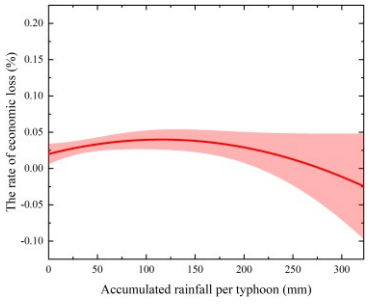
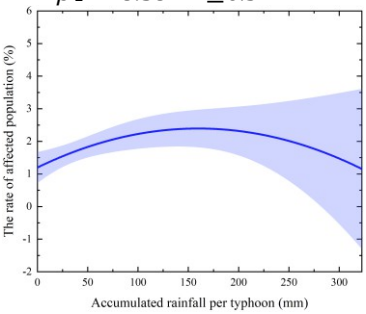
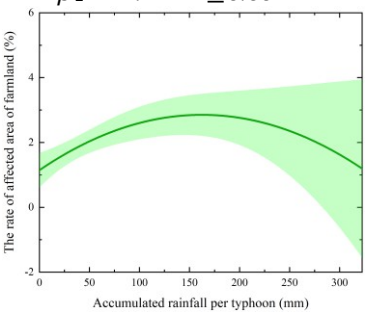
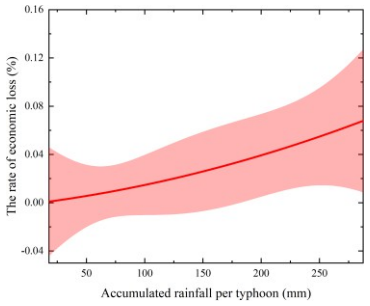
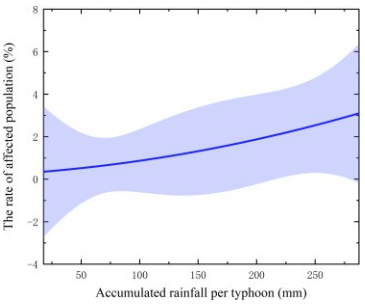
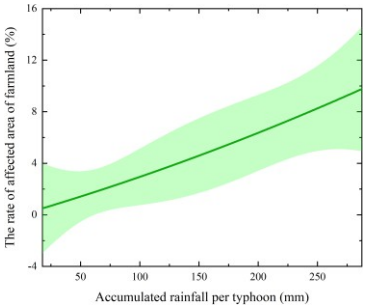
The Optimal Vulnerability Functions of Economic, Population, and Crop in Different Provinces

The refined data of all single typhoons transiting the study area from 2003 to 2018 were collected in this study. The accumulated rainfall during the typhoons at various meteorological stations were compiled as the independent variable. For all the 12 provinces, three kinds of disaster-bearing bodies are considered, including economic, population, and crops. The data of the disaster-bearing bodies were normalized by dividing the loss values by the total amount, that is, the area of affected crop is divided by the sown area; the affected population is divided by the year-end population; and the economic loss is divided by the year-end GDP, respectively. Finally, the economic loss rate, the affected population rate, and the affected crop rate were obtained and then considered as the dependent variables in this study.

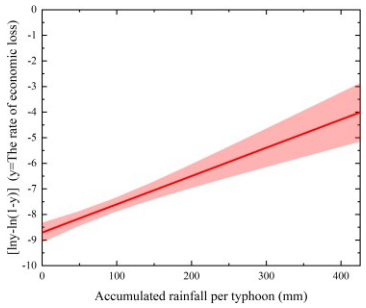
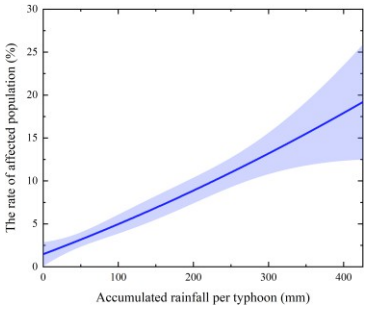
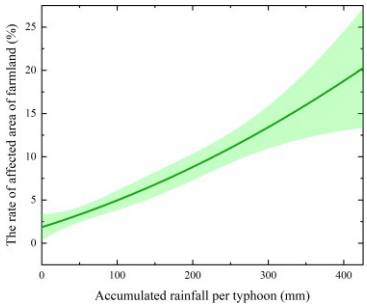
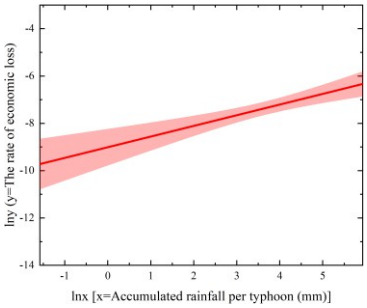
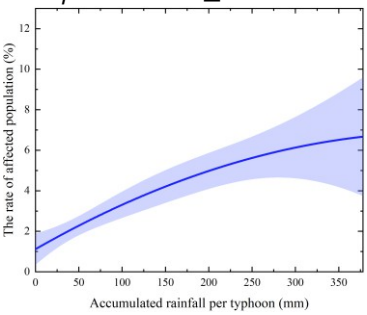
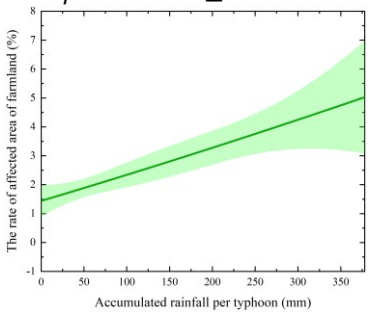
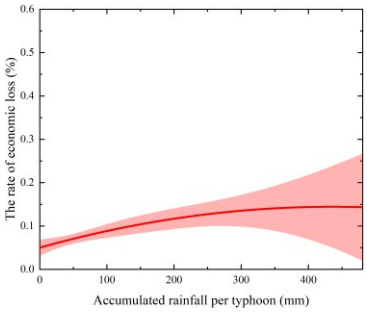
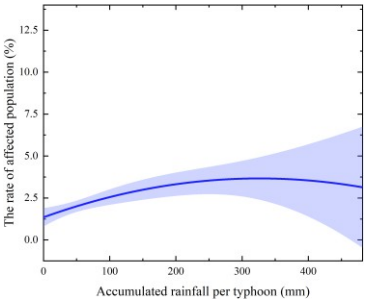
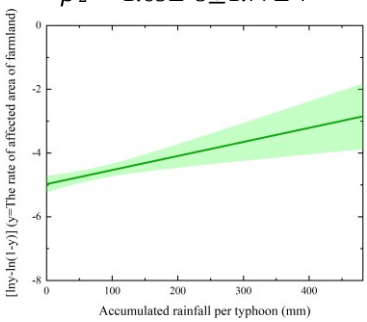
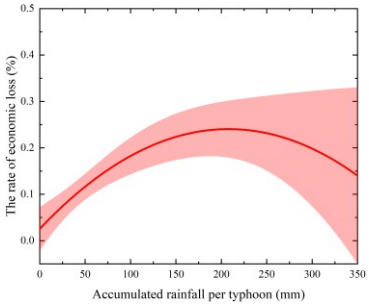
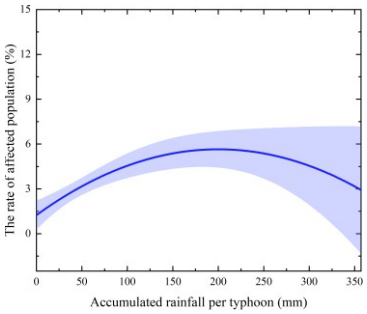
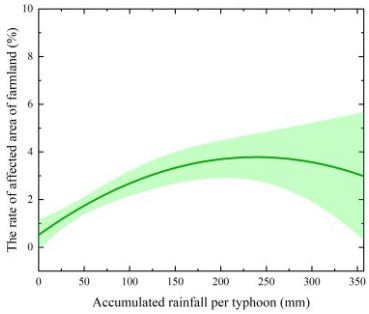
The provincial data were fitted to alternative vulnerability models introduced in Methodology section and the optimal vulnerability function was finally selected by comparing the *R*-squared values of different curves. A 95% confidence interval was plotted for each optimal vulnerability curve to depict the potential range of $f(h)$ at a 95% confidence level. The aggregated results of all the optimal vulnerability functions are presented in Table 2 below.

Table 2. The Vulnerability Functions of Economic, Population, and Crop in Different Provinces

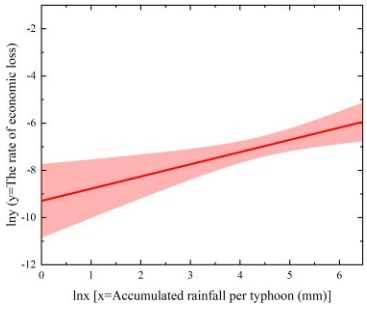
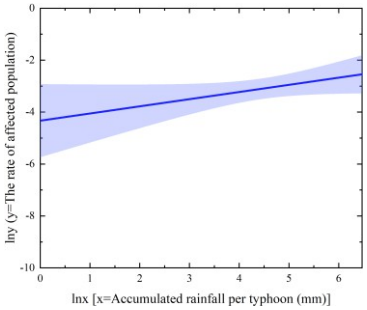
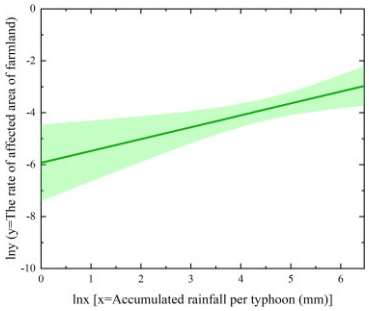
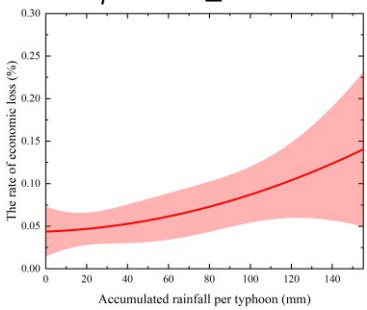
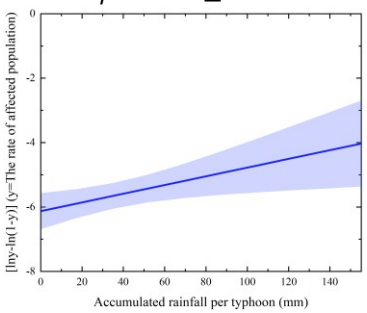
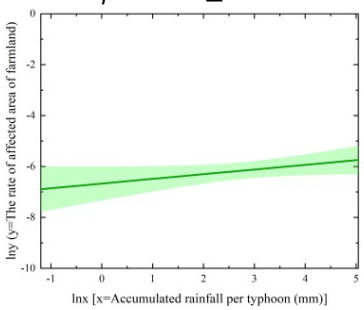
Prov.	Economic Loss Rate	Affected Population Rate	Affected Crop Rate
Jilin	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -10.42±0.77; β_1 = 0.94±0.20	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -7.13±1.08; β_1 = 0.76±0.28	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -6.51±0.64; β_1 = 0.54±0.17
Liaoning	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -9.08±0.90; β_1 = 0.34±0.23	 $Lr = \beta_0 + \beta_1 \ln(h)$, β_0 = 1.21E-2±9.74E-3; β_1 = 2.70E-3±2.51E-3	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 1.84E-2±6.30E-3; β_1 = 1.55E-4±2.00E-4; β_2 = -5.56E-7±1.17E-6

Prov.	Economic Loss Rate	Affected Population Rate	Affected Crop Rate
Hebei			
	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 1.66E-3±8.37E-4; β_1 = -3.83E-5±2.20E-5; β_2 = 2.58E-7±1.09E-7	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 2.32E-2±8.86E-3; β_1 = -4.18E-4±2.32E-4; β_2 = 2.80E-6±1.16E-6	$\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -5.23±0.48; β_1 = -0.11±0.13
Shandong			
	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 =2.80E-4±2.56E-4; β_1 = 1.23E-5±7.73E-6; β_2 = -3.82E-8±3.80E-8	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 8.19E-3±4.70E-3; β_1 = 2.45E-4±1.42E-4; β_2 = -8.58E-7±6.97E-7	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 7.43E-3±4.49E-3; β_1 = 2.23E-4±1.36E-4; β_2 = -7.27E-7±6.66E-7
Jiangsu			
	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 1.99E-4±7.13E-5; β_1 = 3.46E-6±1.70E-6; β_2 = -1.50E-8±7.52E-9	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 1.20E-2±2.40E-3; β_1 = 1.50E-4±5.81E-5; β_2 = -4.69E-7±2.53E-7	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 1.14E-2±2.67E-3; β_1 = 2.10E-4±6.46E-5; β_2 = -6.48E-7±2.82E-7
Shanghai			
	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = -1.29E-5±2.78E-4; β_1 = 1.17E-6±4.68E-6; β_2 = 4.31E-9±1.49E-8	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 2.78E-3±1.86E-2; β_1 = 3.78E-5±3.07E-4; β_2 = 2.10E-7±9.47E-7	$Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 1.90E-4±2.19E-2; β_1 = 2.68E-4±3.92E-4; β_2 = 2.47E-7±1.26E-6

WiP Paper – Impact-driven Emergency and Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Prov.	Economic Loss Rate	Affected Population Rate	Affected Crop Rate
Zhejiang	 $\ln\left(\frac{Lr}{1-Lr}\right) = \beta_0 + \beta_1 h,$ β_0 = -8.71±0.19; β_1 = 1.10E-2±1.64E-3	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 = 1.47E-2±7.04E-3; β_1 = 3.30E-4±1.36E-4; β_2 = 2.03E-7±4.41E-7	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 = 1.85E-2±7.55E-3; β_1 = 2.73E-4±1.43E-4; β_2 = 3.75E-7±4.59E-7
Fujian	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h),$ β_0 = -9.01±0.39; β_1 = 0.45±0.10	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 = 1.12E-2±3.91E-3; β_1 = 2.45E-4±7.90E-5; β_2 = -2.59E-7±2.63E-7	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 = 1.44E-2±2.77E-3; β_1 = 8.85E-5±5.36E-5; β_2 = 1.69E-8±1.77E-7
Guangdong	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 = 5.01E-4±9.50E-5; β_1 = 4.33E-6±1.82E-6; β_2 = -4.98E-9±5.61E-9	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 = 1.35E-2±2.78E-3; β_1 = 1.42E-4±5.32E-5; β_2 = -2.18E-7±1.63E-7	 $\ln\left(\frac{Lr}{1-Lr}\right) = \beta_0 + \beta_1 h,$ β_0 = -4.97±0.13; β_1 = 4.41E-3±1.25E-3
Guangxi	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 = 2.54E-4±2.39E-4; β_1 = 2.07E-5±5.10E-6; β_2 = -4.97E-8±1.85E-8	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 = 1.24E-2±4.87E-3; β_1 = 4.41E-4±1.06E-4; β_2 = -1.10E-6±3.89E-7	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2,$ β_0 =5.17E-3±3.25E-3; β_1 = 2.73E-4±6.87E-5; β_2 = -5.73E-7±2.48E-7

WiP Paper – Impact-driven Emergency and Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Prov.	Economic Loss Rate	Affected Population Rate	Affected Crop Rate
Hainan	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -9.29±0.79; β_1 = 0.52±0.17	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -4.33±0.71; β_1 = 0.28±0.15	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -5.93±0.74; β_1 = 0.46±0.16
Yunnan	 $Lr = \beta_0 + \beta_1 h + \beta_2 h^2$, β_0 = 4.37E-4±1.47E-4; β_1 = 9.17E-7±6.42E-6; β_2 = 3.43E-8±5.02E-8	 $\ln\left(\frac{Lr}{1-Lr}\right) = \beta_0 + \beta_1 h$, β_0 = -6.13±0.28; β_1 = 1.35E-2±5.42E-3	 $\ln(Lr) = \beta_0 + \beta_1 \ln(h)$, β_0 = -6.67±0.33; β_1 = 0.18±0.10

It can be seen from Table 2 that the optimal vulnerability functions for all the three disaster-bearing bodies mainly concentrate on two categories: polynomial model and logarithmic-logarithmic model, accounting for 64% and 25% of the total, respectively. It means that the polynomial model and logarithmic-logarithmic model can more effectively capture the nonlinear relationship between the disaster intensity and the loss rate of the disaster-bearing bodies, thereby showing better fitting performance.

It is found that the optimal vulnerability function forms for the three disaster-bearing bodies may vary even in the same province. For instance, the vulnerability function types for the three disaster-bearing bodies are the same in Jilin, Shandong, Jiangsu, Shanghai, Guangxi, and Hainan, while differ from one another in all the other provinces. Therefore, when studying vulnerability of different disaster-bearing bodies in a specific region, the optimal vulnerability function type should be chosen based on the actual fitting performance, and it is not advisable to simply use a single type of vulnerability function for analysis.

For the provinces of Jilin, Shandong, Jiangsu, Shanghai, Guangxi, and Hainan, despite that all the three vulnerability functions have the consistent form, the parameters from the curve fitting are not the same, indicating that the loss rate variabilities of different disaster-bearing bodies are not the same. Taking Hainan Province as an example, the vulnerability function forms of economic, population, and crop conform to logarithmic-logarithmic model while the parameter β_1 in the models are different, showing the following rules: β_1 of the economic loss rate model > β_1 of the affected crop rate model > β_1 of the affected population rate model. The results suggest that changes in the typhoon intensity have different impact mode and degree on different disaster-bearing bodies.

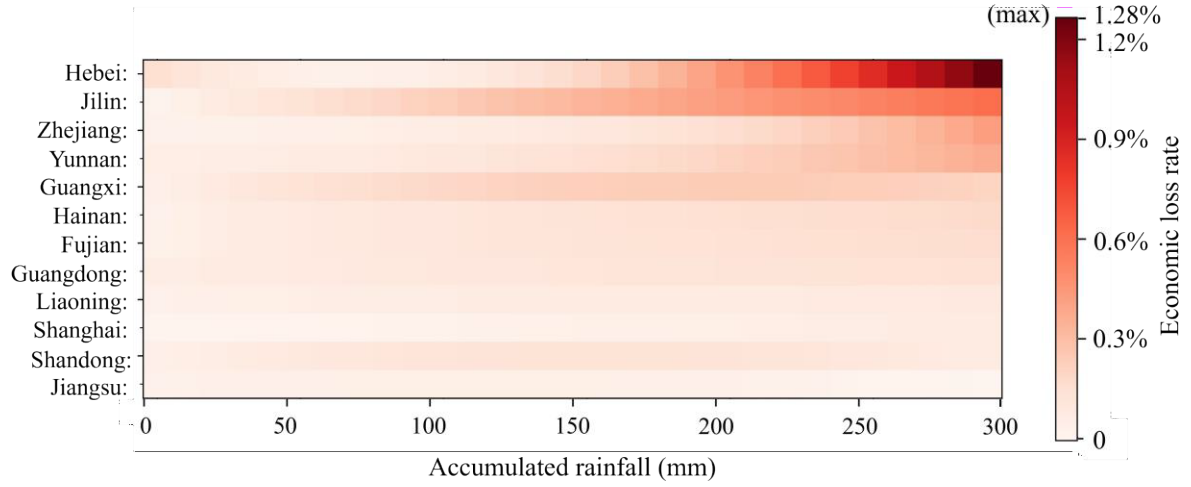
Comparative analysis of the vulnerability functions of the same disaster-bearing body in different provinces reveals that the styles of vulnerability functions usually vary for different regions, even when the optimal fitting vulnerability function types are consistent. For example, Jilin, Liaoning, Fujian, and Hainan share the same form of vulnerability function while the key parameters in the functions are diverse from each other, with β_1 in the model in Jilin > β_1 in the model in Hainan > β_1 in the model in Fujian > β_1 in the model in Liaoning. This reflects that the resilience of different provinces to typhoon disasters could be different.

The Vulnerability Heterogeneity of Economic, Population, and Crop in Different Provinces

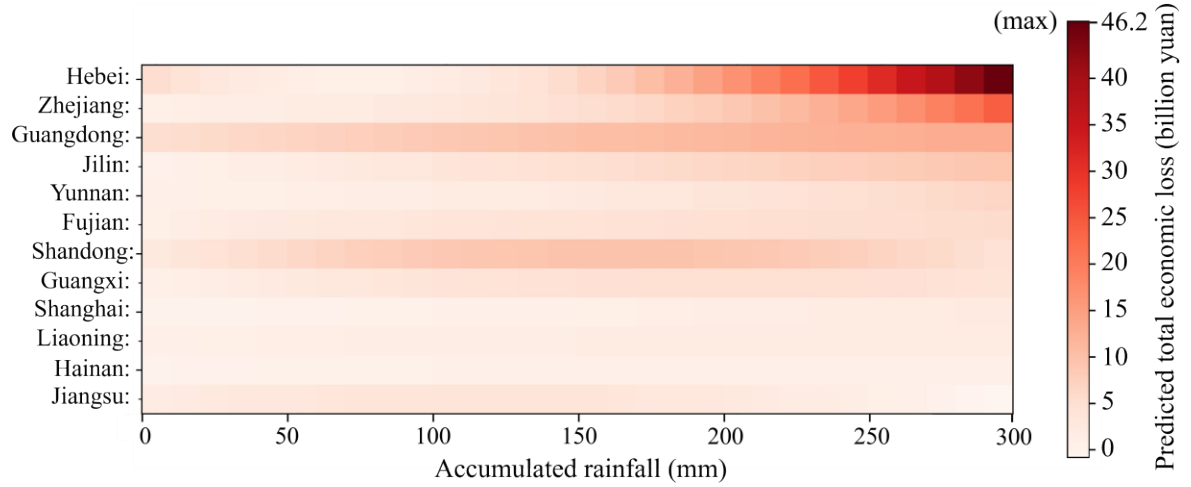
Preliminary conclusions can be drawn by comparing the size of key parameters when the same vulnerability

function form is applied in different disaster-bearing bodies in different provinces, but it is difficult to carry out comparative studies intuitively for different vulnerability function forms. To solve this problem, further visual analysis of the loss rates and predicted loss values (based on the 2018 data) of the three disaster-bearing bodies in 12 provinces under the same disaster intensity is conducted (Figure 3), so as to see the vulnerability heterogeneity of economic, population, and crop in different provinces more directly. It can be seen from Figure 3 that the loss rates and predicted loss values of all the three disaster-bearing bodies generally increase with the increase of the accumulated rainfall during the typhoon, suggesting that more severe disasters can bring greater losses, which is consistent with people's perception. More specifically, there are the following important findings.

As to the heterogeneity of economic loss rate in extreme typhoon-rainfall scenarios in different provinces, it can be seen from Figure 3(a) that Hebei, Jilin, Zhejiang, and Yunnan top the list. It can be seen from Table 1 above that among these provinces, Hebei, Jilin, and Yunnan have the fewest typhoons per year on average. It can be reasonable to speculate that these provinces may be less prepared for typhoons since typhoons are not frequent in history, resulting in greater impacts when typhoons occur. It reminds us that relevant departments should deal with extreme disasters such as typhoons that occur infrequently. Figure 3(b) shows that Hebei, Zhejiang, Guangdong, Jilin, and Yunnan top the list of the predicted total economic loss in extreme typhoon-rainfall scenarios. The ranking of the predicted total economic loss does not change much in general, while of course the ranking of provinces with large Gross Regional Product, such as Guangdong, will rise.



(a) Economic loss rate under the same scenario of typhoon-induced rainfall

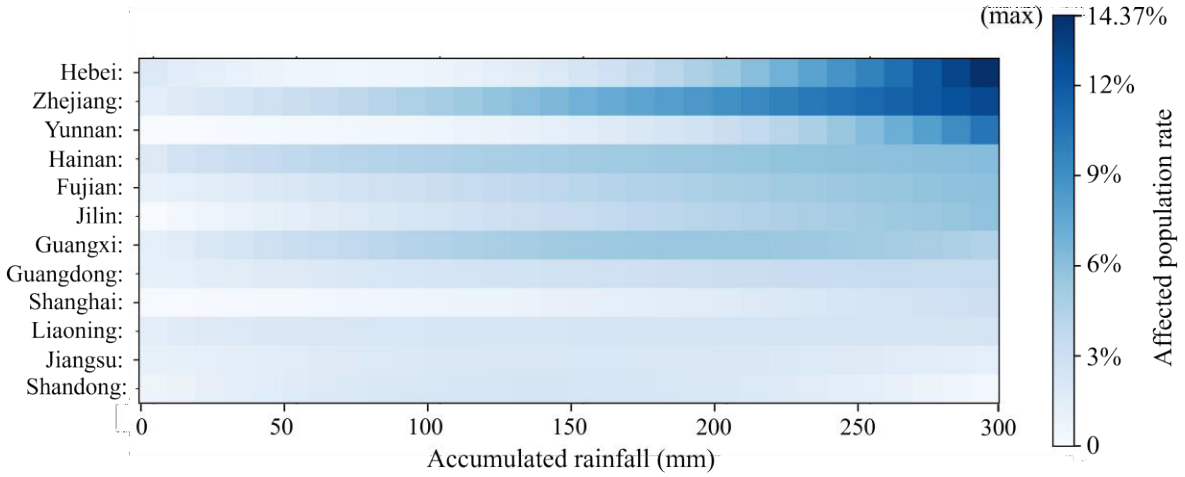


(b) Predicted total economic loss under the same scenario of typhoon-induced rainfall

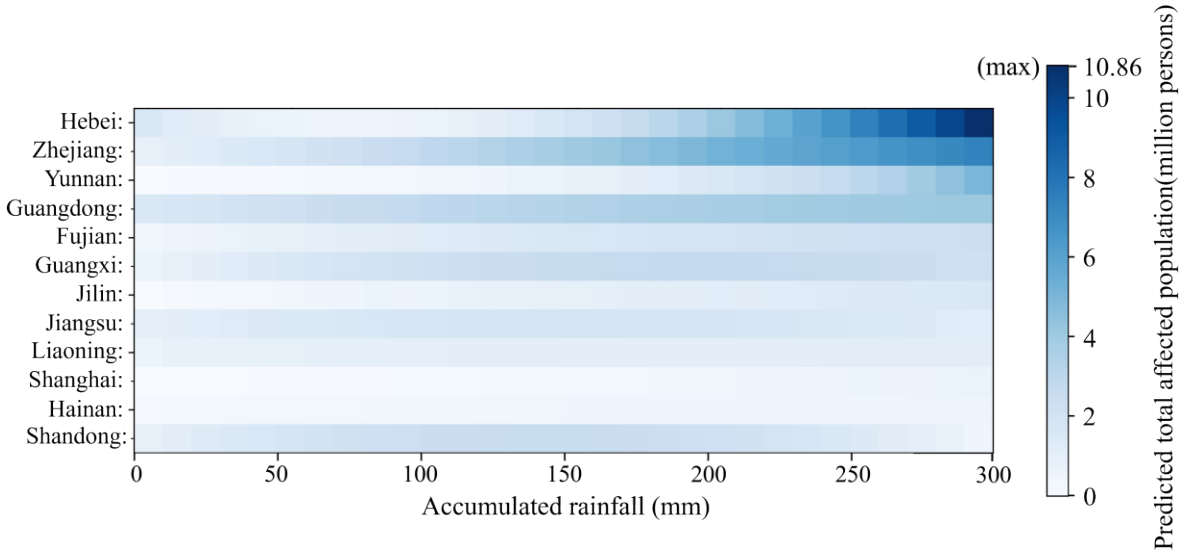
Figure 3. Heterogeneity of Economic Loss in Different Provinces

As to the heterogeneity of affected population rate in extreme typhoon-rainfall scenarios in different provinces, it can be seen from Figure 4(a) that Hebei, Zhejiang, Yunnan, Hainan, and Jilin top the list. This finding is highly

consistent with the finding related to economic loss, suggesting a high similarity degree between the vulnerability of the affected population and the vulnerability of the economic system in these provinces, most of which have experienced less typhoon damage in the history. It can be seen from Figure 4(b) that Hebei, Zhejiang, Yunnan, Guangdong, Fujian top the list of predicted total affected population in extreme typhoon-rainfall scenarios in different provinces. Considering that Guangdong has 113.46 million population and Fujian has 39.41 million population, which are significantly larger than the population in Hainan (9.34 million population) and Jilin (27.04 million population), this result is not difficult to understand.



(a) Affected population rate under the same scenario of typhoon-induced rainfall

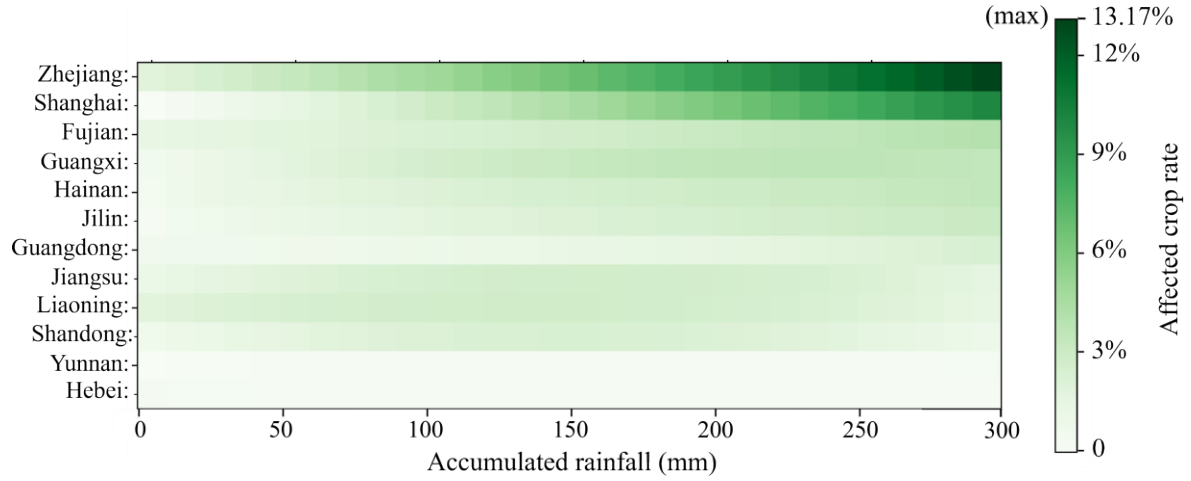


(b) Predicted total affected population under the same scenario of typhoon-induced rainfall

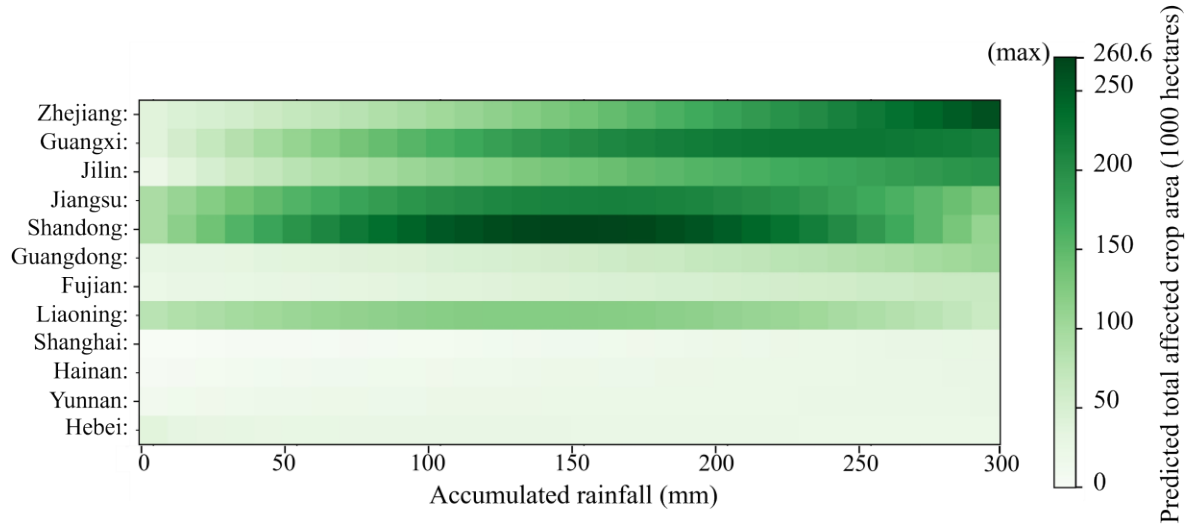
Figure 4. Heterogeneity of Affected Population in Different Provinces

As to the heterogeneity of affected crop rate in extreme typhoon-rainfall scenarios in different provinces, it can be seen from Figure 5(a) that Zhejiang and Shanghai top the list, with obviously larger value of affected crop rate. However, the pattern in Figure 5(b) is a little more complex. Firstly, the predicted total affected crop area in Shanghai is very small, only ranking the fourth from bottom. The main reason for this phenomenon is that Shanghai, as a modern metropolis, has almost no agricultural production. The total sown area in Shanghai is only 282.3×10^3 hectares, and the sown area is even less than 1/10 of that of some provinces. Despite that the affected crop rate in extreme typhoon-rainfall scenarios in Shanghai is huge, the overall crop loss is small. Secondly, due to the regional differences in the crop area, many provinces, although the crop loss rate is small, the total crop loss under extreme typhoon-rainfall scenarios could be large. These provinces include Guangxi (4th in loss rate; 2nd in loss), Jilin (6th in loss rate; 3rd in loss), Jiangsu (8th in loss rate; 4th in loss), and Shandong (10th in loss rate; 5th in loss). These provinces have a common feature that they are traditionally large agricultural provinces, with

a large area of cultivation. Therefore, only the affected crop rate cannot accurately reflect the overall situation of agricultural damage while total affected crop area could be a more accurate indicator. Thirdly, it is found that Zhejiang Province is a typical province with severe agricultural damage under typhoon disasters. Not only the crop loss rate is high in Zhejiang, but also the total crop loss value is extremely large.



(a) Affected crop rate under the same scenario of typhoon-induced rainfall



(b) Predicted total affected crop area under the same scenario of typhoon-induced rainfall

Figure 5. Heterogeneity of Affected Crop Area in Different Provinces

LIMITATIONS AND FUTURE WORK

Limitation in The Selection of Disaster Indicator

When a typhoon passes through, there are often multiple dimensions of disaster indicators. As a preliminary study, only cumulative rainfall during the typhoon transit was selected as the indicator of disaster intensity due to the accessibility of the data in this study, and other indicators were not considered temporarily. In the further study, we plan to furtherly investigate the applicability of other disaster indicators as the independent variable in TC vulnerability function construction for a more accurate and comprehensive typhoon risk assessment.

Limitation in The Selection of Potential TC Vulnerability Functions

Appropriate choice of the potential vulnerability curve forms is the basis for developing effective vulnerability curves. In this study, the six potential vulnerability functions suggested by the work by Jiang et al. (2023) were

mainly referred to. The authors plan to test more forms of vulnerability functions, and hope that more regionally applicable vulnerability curves might be found in near future.

Limitation in Limited Number of Data Samples in Domain of High Disaster Intensity

Higher disaster intensity typically results in greater losses. However, fewer data points from infrequent high-intensity typhoons can disrupt curve accuracy, which causes the problem that the vulnerability curve does not increase monotonously. The authors plan to accumulate typhoon data over a longer period of time, thereby increasing the number of high-intensity-value points to obtain more accurate fitting results.

CONCLUSIONS

Tropical cyclones are one of the natural disasters with high frequency, strong destructive power, and widespread global impact. The southeast coastal area of China is the most economically developed and densely populated area in China, and is also a typhoon-prone area. Therefore, conducting a disaster-bearing body vulnerability assessment in the coastal areas of China is imperative. Based on the collected disaster data in 12 provincial-level coastal regions in China from 2003 to 2018, this study selected accumulated rainfall during the typhoons as the causative factor. The direct economic loss rate, affected population rate, and affected crop rate were selected as the indexes of the disaster outcome. Through the construction, fitting, and comparation of six alternative regression-based vulnerability functions, the optimal vulnerability functions for three disaster-bearing bodies in 12 provinces in were identified.

It was found in this study that the optimal vulnerability curves for all the disaster-bearing bodies are mainly concentrated in polynomial model and logarithmic-logarithmic model, accounting for 64% and 25% of the total, respectively, which shows that these two models could be the most suitable choice for the application of vulnerability assessment in the study area. It was also found that the constructed vulnerability curves generally exhibit a monotonically increasing trend, indicating that losses continuously increase with the increase of cumulative typhoon rainfall. As to the heterogeneity in different provinces, it was found that Hebei, Jilin, Zhejiang, Yunnan, and Hainan top the list of economic loss rate and affected population rate, while most of them have the fewest typhoons per year on average. It indicates that the economic loss and the population suffering condition may be severer in the provinces where typhoons are not common relatively, probably since these provinces were less well prepared for typhoons. Also, it was found that the provincial differences in crop loss rates and crop loss values are very large. The loss rate in large agricultural provinces is not high, but the absolute loss value could be very large, reflecting that crop losses are very dependent on the size of agricultural acreage, rather than solely determined by the loss rates.

As a preliminary study, there are still some limitations in this paper. To the best of the authors' knowledge, this paper gives a first attempt in exploring the optimal vulnerability functions for coastal provinces in China. The relevant results from this paper are of great significance for understanding the spatial heterogeneity of typhoon disasters and disaster victims. This paper could also provide valuable references for government departments in implementing accurate typhoon risk assessment, and improving disaster prevention and response capabilities.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (Grant No. 72304038), Guangdong Basic and Applied Basic Research Foundation (Grant No. 2022A1515110049), Young Innovative Talents Project from Department of Education of Guangdong Province (Grant No. 2022KQNCX157).

DATA AVAILABILITY STATEMENT

All the data used in this article are provided with references and links at the manuscript.

REFERENCES

Academy of Disaster Reduction and Emergency Management, National Disaster Reduction Centre of China, International Federation of Red Cross and Red Crescent Societies (IFRC), Beijing Normal University. (2023). 2022 Global natural disaster assessment report. https://www.preventionweb.net/publication/2022-global-natural-disaster-assessment-report

Acosta, T.J.S., Galisim, J.J., Tan, L.R., Hernandez, J.Y. (2018). Development of empirical wind vulnerability curves of school buildings damaged by the 2016 Typhoon Nina. *Procedia Engineering*, *212*, 395-402. https://doi.org/10.1016/j.proeng.2018.01.051

Barnett, J. (2020). Global environmental change II: Political economies of vulnerability to climate change. *Progress in Human Geography*, *44*(6), 1172-1184. https://doi.org/10.1177/0309132519898254

China Meteorological Administration. (2004-2019). Yearbook of Meteorological Disasters in China. https://data.cnki.net/statisticalData/index?ky=%E4%B8%AD%E5%9B%BD%E6%B0%94%E8%B1%A1%E7%81%BE%E5%AE%B3%E5%B9%B4%E9%89%B4

Colombi, M., Borzi, B., Crowley, H., Onida, M., Meroni, F., & Pinho, R. (2008). Deriving vulnerability curves using Italian earthquake damage data. *Bulletin of Earthquake Engineering*, *6*, 485-504. https://doi.org/10.1007/s10518-008-9073-6

Godfrey, A., Ciurean, R.L., van Westen, C.J., Kingma, N.C., & Glade, T. (2015). Assessing vulnerability of buildings to hydro-meteorological hazards using an expert based approach - An application in Nehoiu Valley, Romania. *International Journal of Disaster Risk Reduction*, *13*, 229-241. https://doi.org/10.1016/j.ijdrr.2015.06.001

Jiang, X.Y., Ma, X.Y., & Yang, L.J. (2023). Comprehensive comparative study on the construction method of flood vulnerability and fragility curves under the framework of regression analysis. *Shuili Xuebao*, *54*(2), 184-198. https://doi.org/10.13243/j.cnki.slxb.20220307 [in Chinese]

Lee, T.C., Knutson, T.R., Nakaegawa, T., Ying, M., & Cha, E.J. (2020). Third assessment on impacts of climate change on tropical cyclones in the Typhoon Committee Region - Part I: Observed changes, detection and attribution. *Tropical Cyclone Research and Review*, *9*(1), 1-22. https://doi.org/10.1016/j.tcrr.2020.03.001

Lu, X. Q., Yu, H., Ying, M., Zhao, B.K., Zhang, S., Lin, L.M., Bai, L.N., & Wan, R.J. (2021). Western North Pacific tropical cyclone database created by the China Meteorological Administration. *Advances in Atmospheric Sciences*, *38*(4), 690−699. https://doi.org/10.1007/s00376-020-0211-7

Ministry of Natural Resources of the People's Republic of China. (2017). China Marine Statistical Yearbook. https://cnki.nbsti.net/CSYDMirror/trade/Yearbook/Single/N2019060088?z=Z007

Monteleone, B., Borzì I., Bonaccorso, B., & Martina, M. (2023). Quantifying crop vulnerability to weather-related extreme events and climate change through vulnerability curves. Natural Hazards, 116, 2761–2796. https://doi.org/10.1007/s11069-022-05791-0

National Bureau of Statistics of China. (2004-2019). China Statistical Yearbook. https://data.cnki.net/yearBook/single?nav=%E7%BB%9F%E8%AE%A1%E5%B9%B4%E9%89%B4&id=N2023110024

National Meteorological Information Center. (2023). Typhoon. https://data.cma.cn/article/getLeft/id/326/keyIndex/3.html [in Chinese]

Papathoma-Köhle, M., Gems, B., Sturm, M., & Fuchs, S. (2017). Matrices, curves and indicators: A review of approaches to assess physical vulnerability to debris flows. Earth-Science Reviews, 171, 272-288.

Pinelli, J.P., Pita, G., Gurley, K., Torkian, B., Hamid, S., & Subramanian, C. (2011). Damage characterization: application to Florida public hurricane loss model. *American Society of Civil Engineers*, *12*(4), 190. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000051

Shifa, M., Gordon, D., Leibbrandt, M., & Zhang, M. (2022). Socioeconomic-related inequalities in COVID-19 vulnerability in South Africa. *International Journal of Environmental Research and Public Health*, *19*(17), 10480. https://doi.org/10.3390/ijerph191710480

Vamvatsikos, D., Kouris, L. A., Panagopoulos, G., Kappos, A. J., Nigro, E., Rossetto, T., & Stathopoulos, T. (2010). Structural vulnerability assessment under natural hazards: a review. In COST ACTION C (Vol. 26, pp. 17-18).

Vickery P.J., Lin, J., Skerlj, P.F., Twisdale, L.A., & Huang, K. (2006). HAZUS-MH hurricane model methodology I: hurricane hazard, terrain, and wind load modeling. *Natural Hazards Review*, *7*(2), 82. https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(82)

Vickery P.J., Skerlj, P.F., Lin, J., Twisdale, L.A., Young, M.A., & Lavelle F.M. (2006). HAZUS-MH hurricane model methodology II: damage and loss estimation. *Natural Hazards Review*, *7*(2), 94. https://doi.org/10.1061/(ASCE)1527-6988(2006)7:2(94)

Wan, C.C., Yan, Y.F., Shen, L.C., Liu, J.L., Lai, X.X., Qian, W., Nie, J., & Wen, J.H. (2023). Damage analysis of retired typhoons in mainland China from 2009 to 2019. *Natural Hazards*, *116*, 3225-3242. https://doi.org/10.1007/s11069-022-05804-y

Wang, S., & Toumi, R. (2021). Recent migration of tropical cyclones toward coasts. *Science*, *371*(6528), 514-+. https://doi.org/10.1126/science.abb9038

White, G.O., Boddewyn, J.J., Rajwani, T., & Hemphill, T.A. (2018). Regulator vulnerabilities to political pressures and political tie intensity: The moderating effects of regulatory and political distance. *Management International Review*, *58*, 743-769. https://doi.org/10.1007/s11575-018-0351-5

World Meteorological Organization (WMO). (2022). Tropical cyclones. https://public-old.wmo.int/en/our-mandate/focus-areas/natural-hazards-and-disaster-risk-reduction/tropical-cyclones

Xu, X.L., Sun, D.W., & Guo, T.J. (2015). A systemic analysis of typhoon risk across China, *Natural Hazards*, *77*(1), 461-477. https://doi.org/10.1007/s11069-015-1586-0

Xu, X., Qiao, S.R., Jiang, H.L., & Zhang, T. (2023). Ecosystem vulnerability to extreme climate in coastal areas of China. *Environmental Research Letters*, *18*(12), 124028. https://doi.org/10.1088/1748-9326/ad0860

Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., Zhu, Y., & Chen, D. (2014). An overview of the China Meteorological Administration tropical cyclone database. *Journal of Atmospheric and Oceanic Technology*, *31*, 287-301. https://doi.org/10.1175/JTECH-D-12-00119.1

Zhang, Y., & Wang, H. (2016). A review of research on typhoon storm surge disaster risk assessment. Marine Forecasts, (02), 81-88. https://doi.org/10.11737/j.issn.1003-0239.2016.02.012

Zhong, X.C., Fang, W.H., & Cao, S.J. (2017). Probabilistic component-based Monte Carlo simulation of vulnerability for a typical low-rise rural residential building in coastal China. *Journal of Beijing Normal University (Natural Science)*, *53*(01), 51-59. https://doi.org/10.16360/j.cnki.jbnuns.2017.01.010 [in Chinese]

Zhou, Y., & Wang, J. (2012). A review on development of vulnerability curve of natural disaster. Advances in Earth Science, 27(4), 435-442.

Zhu, J.T., Liu, K., Wang, M., Xu, W., Liu, M.T., Zheng, J.C. (2022). An empirical approach for developing functions for the vulnerability of roads to tropical cyclones. *Transportation Research Part D: Transport and Environment*, *102*, 103136. https://doi.org/10.1016/j.trd.2021.103136

Crisis Response and Management of Public Administrations: Forgotten Tales

Yannic Schulte
University of Wuppertal
Institute for Public Safety and Emergency Management
yschulte@uni-wuppertal.de

Malte Schönefeld
University of Wuppertal
Institute for Public Safety and Emergency Management
schoenefeld@uni-wuppertal.de

Patricia M. Schütte
University of Wuppertal
Institute for Public Safety and Emergency Management
schuette@uni-wuppertal.de

Frank Fiedrich
University of Wuppertal
Institute for Public Safety and Emergency Management
fiedrich@uni-wuppertal.de

ABSTRACT

Crises and disasters have become omnipresent worldwide in recent years. But thinking of managing those critical situations the focus of the public and scientific research in recent years was primarily on blue light organizations, i.e., classic emergency response organizations. This changed with the refugee situation in 2015/2016 and the COVID-19 pandemic, when public administrations and local authorities in particular took on a central role in crisis management. They became the central players in those situations, played an important role at the interface with the local population and were the first point of contact and information for them. This article sheds light on municipal administrations in Germany which have often be overlooked in various stages of crisis management. Drawing on empirical evidence from a German research project, it delves into the extent to which municipalities may have been overlooked or disregarded in certain aspects.

Keywords

Public administrations, Crisis management, Refugee movements, COVID-19, Empirical research

INTRODUCTION

Crises and disasters have become omnipresent worldwide in recent years. Authorities and organizations with safety and security tasks (AOS) implement measures to deal with these events on the basis of their (organizational) structures and processes, and developed and tested operational plans. Thinking of crises and disasters, the focus of the public, but also of scientific research in the last few years, was primarily on blue light organizations, i.e., classic emergency response organizations like the German Federal Agency for Technical Relief or fire brigades. On the basis of an ongoing research project, we would like to draw attention to an actor apart from the classic “blue light organizations” who, in our view, plays an important role in all phases of crisis management: local administrative authorities or municipalities. While "usual" crises such as extreme weather events are particularly challenging for established emergency organizations, the refugee situation 2015/16 and the COVID-19 pandemic are examples of how important crisis-proof and crisis-tested structures are, not only in established emergency management organizations, but also in public administrations. These situations are characterized by the range of tasks that arise for public administrations (on different levels) in the context of crisis management, which makes them the central players in those situations. They played an important role at the interface with the local population and were the first point of contact and information for them. During the refugee situation in 2015/16,

WiP Paper – Impact-driven Emergency and Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

administrations took on tasks in the area of health protection, registration and accommodation of refugees. During the COVID-19 pandemic, administrations had to compensate for staff shortages, while at the same time the number of personnel-intensive tasks, such as contact tracing and vaccine prioritization and distribution, increased. This clearly shows that administrations play a key role in crisis situations. Therefore, it is necessary for them to establish processes for the preparation, management and also recovery phase. Personnel in public administrations needs the necessary training and flexibility for effective crisis management, for which sufficient resources and skills are required even before the crisis situation arises in order to be able to react quickly and flexibly.

This article sheds light on municipal administrations in Germany which have often be overlooked in various stages of crisis management. It will be discussed to which extent local administrations or local authorities may have been forgotten or neglected in some respects in an integrated perspective on crisis management. Based on empirical interview material on the handling of the COVID-19 pandemic and the analysis of media and administrative documents in Germany, we look at how local authorities have positioned themselves in the area of crisis management in this long-lasting situation. We also examine how cooperation between local authorities and other crisis management stakeholders has developed and evolved and why the preparation and follow-up phase of the crisis management lifecycle play an important role in integrated crisis management.

RESEARCH PROJECTS: DATA COLLECTION AND METHODS

The project " *Local governments in crisis mode (KoViK)*", which is funded by the German Research Foundation between 2022 and 2024, deals with the management of the COVID-19 pandemic from a municipal perspective. The project builds on previous research on dealing with the refugee situation in 2015/2016 from the perspective of various actors, including various public administrations. The resulting findings revealed the following:

(1) that administrations had little practice in crisis management at that time, that administration representatives did not see themselves or their organizations as central authorities in crisis management and that crisis management was not a central, but merely an "additional task" alongside the actual administrative work;
(2) that they were not perceived by third parties such as AOS as professional representatives of crisis management;
(3) that administrations lacked suitable structures for the adequate management of an acute crisis, but were instead rather inhibited due to bureaucratic principles and a so-called "iron cage" inherent in bureaucracy, and
(4) that in some places they were preoccupied with addressing their internal organizational challenges rather than effectively managing the situation at hand (Schütte et al., 2022; Schulte et al., 2022).

The KoViK project builds on these findings and pursues three objectives: Firstly, using the SARS-CoV-2 situation as an example, insights are to be gained into practiced approaches to crisis management and crisis communication by municipal administrations. Secondly, those are reflected with regard to established forms and, thirdly, the consequences of pluralistic approaches for the population and other actors are anticipated. The project focuses on cases in North Rhine-Westphalia (NRW).

In order to take as comprehensive a look as possible at the crisis management of local authorities, the study uses a multi-method approach, which can be seen in the following figure (see Figure 1).

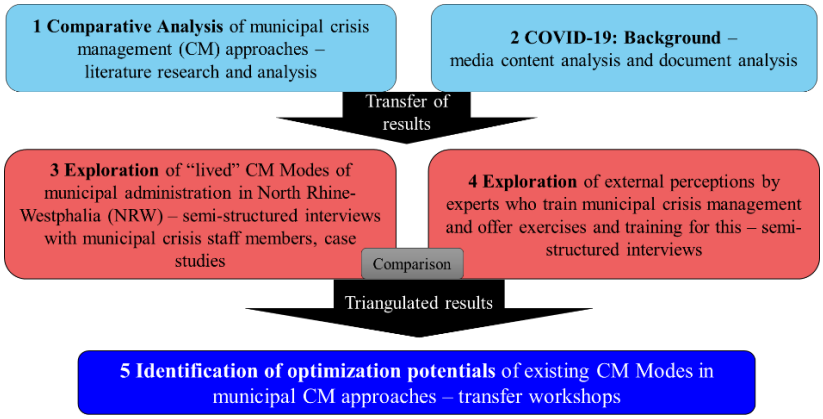


Figure 1. Work structure of the KoViK research project on municipal crisis management (Source: own illustration based on Schönefeld et al. (2023)).

The aim of the first work package (WP) (1) was a comparative analysis of scientific studies that deal with the topic of municipal crisis management. Here, a state of research and the approaches to be derived from it were developed. Parallel to this, a media content analysis on the development of the pandemic in municipalities in NRW was carried out in WP 2. In addition, legal texts in the field of disaster and crises law were analyzed to see how they deal with the concept of crisis management (WP 2). Based on the results of WP 2, 25 semi-structured interviews were conducted: (a) 19 interviews with administrative staff of independent cities and rural districts in NRW who were involved in crisis management activities during COVID-19 pandemic at least, i.e., in crisis teams (self-perception) (WP 3) as well as (b) 6 interviews with experts who provide training in municipal crisis management, design and conduct training, exercises for this purpose and are therefore able to provide an external assessment of municipal crisis management and crisis team work (external perception) (WP 4). The empirical results of WP 1, 2, 3 and 4 will be triangulated and documented in five case studies. In these, the results for the 19 cities and districts are summarized at the level of administrative districts (there are five administrative districts in NRW) in order to ensure a certain degree of anonymity. All of the results will be triangulated and prepared for practice transfer workshops in which practitioners will collaborate to identify opportunities for improvement explore feasible implementation strategies (WP 5). The following sections present the (not yet conclusive) results of WP 1 to 3.

(LOCAL) PUBLIC ADMINISTRATIONS IN CRISIS MANAGEMENT: BRIEF STATE OF THE ART

Global critical situations also affect the local level. This was the case for the COVID-19 pandemic as well. For more than two years, public administrations at all federal levels in Germany operated in a continuous crisis mode. Municipalities, as the "lowest main administrative level(s) responsible for the entire local administration" (Ehl & Wendekamm, 2013, p. 133, own translation) were facing major challenges, because it was their mission to implement decisions made at federal and state level. Municipal governments took over a central part of coping with consequences of the COVID-19 pandemic within their frameworks of local crisis management, while their daily tasks had to be continued. Before that, local and intermediary (here: *Landkreise/Regierungsbezirke*) actors typically gained relatively low attention by the public, media, research, etc. in crises and disasters, compared to “blue light organizations” or the administrative counterparts on the national and federal level (Benton, 2002; Schneider & Park, 1989), although it does not match with the high importance of local and regional emergency management. Major exceptions are events such as the storm surge in Hamburg in 1962, when Helmut Schmidt (then Hamburg's police senator and later German Chancellor) implemented a rapid but legally difficult rescue operation.

At least until 2020, it can be assumed that the state of the art in (administrative) science – especially in the German-language literature – was deficient. Research revealed only a few studies on the crisis management of public administrations in relation to social crises, as the following quote makes clear:

"From an administrative science perspective, disasters and manifest crises are rarely studied. This is particularly surprising because public authorities play a central role in coping with disasters" (Hustedt, 2019, p. 183, own translation).

In recent years, overviews of administrative science on administration in Germany have focused on crises, the associated challenges for German administrations, and the resulting (potential) changes and reforms (e.g., digitalization of administrations) have been discussed, but no mention has been made of specific crisis management approaches by administrations for such social crises, forms of handling such as administrative staff or aspects of crisis communication (e.g., Bauer & Grande, 2018; Bogumil et al., 2016; Bogumil & Jann, 2020; Kuhlmann & Schwab, 2017; Möltgen-Sicking & Winter, 2018; Ziekow, 2018 cited after Schönefeld et al., 2023). One possible reason for this could be that public administrations were less likely to appear as "the" central crisis managers in earlier times. Usually, various public safety and emergency response organizations such as the police and fire department were the focus of social and media attention when it came to dealing with critical situations (Schönefeld et al., 2023). However, this changed with the refugee movements in 2015/2016 and the COVID-19 pandemic. During the pandemic, studies on public administrations – especially local administrations – focused primarily on, e.g., the assessment of measures (by administrative staff and citizens) and the satisfaction of the population with municipal crisis management (e.g. Eckhard & Lenz, 2020; Otto, 2021; Pöhler et al., 2020). In addition, approaches to learning from critical phases were also examined (e.g. Aden, 2023; C. Bauer & Egger, 2023; Bauer et al., 2022). Nevertheless, there were mostly rather superficial considerations of the crisis management of public administrations (Neumann, 2020). Studies examining the distinctions and interconnections among crisis management at various governmental levels – federal, state, district, and local – have only begun to emerge gradually, roughly coinciding with the progression of the pandemic (e.g., Kuhlmann & Franzke, 2022; Lenz, 2023).

When the KoViK was applied for in 2021, the scientific literature on crisis management in (German) local authorities was comparatively "thin" and pointed to more research gaps and associated questions than findings and answers. Against this background, corresponding media content and document analyses (of legal texts and other legal foundations) were carried out in order to create the basis for an understanding of the image of crisis management in local government that is conveyed. The following passages take a closer look at the findings.

(LOCAL) PUBLIC ADMINISTRATIONS: FORGOTTEN TALES IN DOCUMENTS AND MEDIA

Media Content Analysis

As already mentioned, (local) public administrations are an important stakeholder in crisis management. During COVID-19, they were a key player when it came to developing or implementing rules in relation to pandemic management. They were and are an important link between the population (as those affected by the pandemic and the chosen crisis management measures) and the higher crisis management levels on district, state and federal levels. Based on this, the question arises as to whether the role those local administrations have taken on in crisis management has also been reflected in media coverage. This analysis was predicated on the assumption that experiences in complex societies are predominantly mediated through media channels (Sackmann, 2020, p. 354). As the project focuses on one region in Germany, regional newspapers were chosen accordingly. A total of seven newspapers were scrutinized by using a search strategy incorporating terms related to the pandemic, municipal administrations and crisis management (actors) as shown here and translated from German by the authors ((Corona OR COVID-19 OR SARS-CoV-2 OR pandemic) AND (municipality OR public health department OR administration) AND (crisis*)). This strategy was designed to capture articles published from the onset of 2020 through the early stages of 2023, specifically selecting the first complete week of each quarter as sampling points. This temporal delineation was chosen to ensure a comprehensive overview of the media landscape over the specified period. For initial screening of the articles four inclusion and 19 exclusion criteria were formulated, grounded in qualitative theses on crisis management and municipal administration. Two of the 23 criteria are listed here as examples. The inclusion criteria primarily comprised the relevant actors, as well as the organization and processes for crisis management, e.g. “The report is neutral/positive/negative about pandemic management / crisis management (in particular organization/structure/processes of the stakeholders)”. Reports on multiple crises, such as a flood occurring at the same time as the pandemic, were also included. In contrast, articles of less than 300 words were excluded, as were articles that merely reported on the course of the pandemic (e.g., in the form of indices) or, for example, on progress in vaccinating the population. The included articles should explicitly focus on crisis management.

The first search identified a corpus of 674 articles. 128 articles then were removed due to duplicates. From the remaining 546 articles 108 were selected on the basis of the in- and exclusion criteria (see Figure 2).

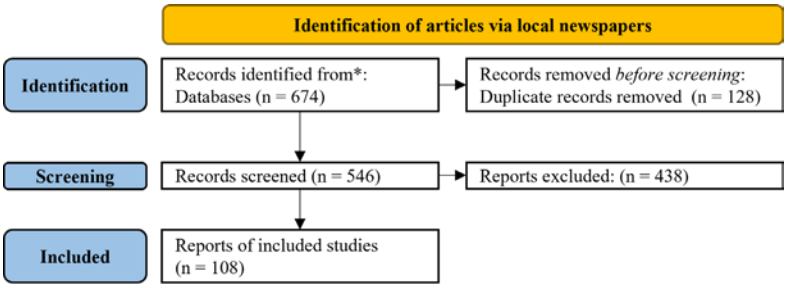


Figure 2. Retrieval of local newspaper articles; own illustration based on Page et al., 2021.

The analysis of the media content has not yet been fully completed. Only initial highlights can be presented here. Using a qualitative thematic synthesis (Thomas et al., 2017) for the analysis illuminates several predominant themes in relation to crisis management. These included the (differences in) municipal handling of the COVID-19 pandemic and the development of case numbers and vaccine availability, challenges for the local citizens like closing of schools, and financial measures implemented by local governments, e.g., for small businesses. A significant emphasis was placed on challenges such as staffing shortages and the imperative for digitalization within municipal crisis management operations. Conversely, less attention was afforded to the structures and processes underpinning crisis management efforts. In summary, explicit references to crisis management were infrequently encountered within the titles or texts of the analyzed articles, suggesting a more implicit engagement with the topic within media discourses. The emphasis in the analyzed media content predominantly lays on the immediate visible aspects and impacts of the crisis, the reactive measures undertaken and the challenges faced, rather than on a detailed examination of strategic or procedural aspects of crisis management. The complexities

of crisis management and more comprehensive public understanding of the roles of municipal administrations in such critical situations remain in the background.

Document analysis of legal background of CM in Germany

A further step in the project presented is the analysis of documents relating to public administrations and their crisis management. The document analysis was intended to examine how the actor and its crisis management is anchored and structured in terms of administrative law. For this purpose, the 16 disaster control laws on state level – because there are no "crisis management laws" or anything comparable – in Germany were analyzed using a qualitative content analysis (Kuckartz & Rädiker, 2022) and compiled in a synopsis. It focused on the statutory frameworks across the federal states of Germany, specifically examining their references to municipal crisis management. This investigation looked into the legal texts to identify the regulations regarding the requirements, structures, processes, and the provisions for personnel and technical equipment outlined therein. Furthermore, the analysis sought to understand how these statutes address the various phases of crisis management in these legislative texts. The primary aim was to identify how these laws articulate the roles and responsibilities at the municipal level, alongside the mechanisms and resources allocated for crisis management and disaster response. The content analysis of the legal documents revealed a few interesting points: A notable observation is that only a few states offer a general definition of crisis management, and even fewer distinguish clearly between the concepts of disaster management and crisis management. This lack of differentiation suggests a conceptual overlap that may affect the clarity and effectiveness of crisis response strategies. Civil protection departments are predominantly designated as the lead actors in these frameworks, a consistent topic across the legislative texts. However, the depth of description regarding the operational teams or staffs for disaster management or crisis management varies significantly, often described in broad terms without delving into the specifics of their formation, roles, or operational guidelines. The general approach to detailing the structure of crisis management efforts underscores a potential gap in providing clear guidance for the execution of these crucial functions at the municipal level. The way in which crisis management work is organized is therefore the responsibility of the actors involved and is therefore left to each local authority. Another critical finding of this analysis is the sporadic mention of the crisis management cycle phases within these laws. Notably, the follow-up phase, essential for learning and improvement post-crisis, is almost never explicitly addressed. This points to a potential oversight in ensuring comprehensive crisis management planning that includes not only immediate response but also the long-term embedding and improvement of the crisis management process. These insights result from the 16 analyzed disaster control laws of the federal states in Germany. The findings highlight the need for clearer definitions, differentiated roles, and detailed procedural guidelines within Germany's disaster management and crisis response statutes. Addressing those gaps through legislative refinement and the incorporation of comprehensive crisis management cycle considerations could significantly enhance the effectiveness and efficiency of municipal crisis management efforts.

However, the document analysis has not yet been completed. Further documents will be included. The authors see the need to look at regulations and decrees at the local level, which can provide additional detail and concrete examples of crisis management practices. For example, the crisis management directive in NRW represents a more detailed formulation of local crisis management, which could provide insights into the operationalization of local crisis management frameworks.

(LOCAL) PUBLIC ADMINISTRATIONS: FORGOTTEN TALES IN EMPIRICS/PRACTICE

This section is about forgotten tales and our findings in the empirical data in terms of interviews that we collected in the KoViK research project. The interviews were semi-structured, i.e., they combined pre-determined questions with the opportunity for the interviewer to explore deeper insights or follow up on responses through additional, impromptu questions. This approach allowed for a flexible conversation while ensuring that specific topics of interest were covered, facilitating both comparability across interviews and in-depth exploration of individual perspectives. The interviews were then transcribed and subsequently coded. Coding of interview content is a systematic process used in qualitative research to organize and categorize the data collected from interviews. This involves identifying patterns, themes, or concepts within the interview transcripts and assigning codes (labels or tags) to segments of text that represent these ideas. Coding serves as a crucial step in interpreting and drawing meaningful conclusions from qualitative data. For this interview material, a deductive-inductive approach was chosen for coding the interviews (Kuckartz & Rädiker, 2022).

We concentrate on four main topics that we consider as particularly relevant.

(1) *The myth of follow-up (respectively evaluation) of actual crisis management practices in municipal administrations*: Follow-up practices are regularly been considered as important in a holistic approach to crisis

management. However, it regularly falls by the wayside or is carried out subordinately. It can be noted that a debriefing is universally desired, but often there is a lack of personnel resources, as one crisis follows another, or the return to day-to-day business leaves no more capacity for in-depth follow-up. Moreover, employees leave the administration faster than expected, taking the relevant knowledge with them:
"What we would like is truly a debriefing of all participants […] Besides the fact that so much has changed, the memories fade. And it's not just me; next year, our department head, who played a major role in handling the pandemic, will no longer be here. And who are you going to work with to process things when the key players involved are no longer around?" (PublicAdministrationInterview(PAI)_15)
In addition to the content, however, the importance of a ceremonial follow-up in the form of thank-you celebrations and small tokens of appreciation was repeatedly mentioned, which are seen as important from an emotional point of view in order to return to normal mode, but also to maintain the willingness to actively support crisis management in future situations. It is the explicit aim of the KoViK research project to gather and systematize the available crisis knowledge in order to be better prepared in the future and to support the often-limited follow-up resources of municipalities.

(2) *Learning on the job and missing exercises*: Exercises play a pivotal role in any civil protection curriculum. In the case of municipal administrations, the relevant training content is how to work in an administrative (not an operational) crisis management staff. Various public institutions in Germany offer these courses (e.g., Federal Office for Civil Protection and Disaster Assistance, Fire Brigade Institute of North Rhine-Westphalia), but resources are limited. On the other hand, interest of municipalities has so far been limited. The size and length of the crisis meant that untrained personnel also had to be deployed in crisis management (and therefore *was learning on the job*).
The task "crisis management" had to be re-learned: *"After the fall of the Cold War wall, I believe, no one thought anymore about disasters, crises, wars. The change has definitely been there for the last two years"* (PAI_3). This departure from the task of crisis management also meant that hardly any resources were available for it: *"The mechanisms that used to exist, and the resources that had been dismantled, are only now slowly being rebuilt."* (PAI_7). This also includes that untrained personnel had to be trained on the job, in the crisis: *"not everyone who was involved there was fully trained. Instead, they had to grow into the task through the crisis"* (PAI_1).

(3) *Empowerment of public servants during crises*: The extent of the COVID-19 pandemic required a large part of the municipal workforce to actively participate in crisis management – in addition to or instead of their actual tasks. Especially in the particularly intense initial phase of the pandemic, municipal staff were occupied with crisis management tasks in multiple shifts over several months. A cliché of working in administration is that of routine and sometimes even boring-monotonous work. The crisis provided staff with the "opportunity" for a different kind of work, across established hierarchical and professional boundaries, and possibly also with a different sense of self-efficacy: *"this mixing of levels was incredibly good for the self-perception of a district administration. Because suddenly people were important whom one might have foolishly thought before were not so important. And they were very important for certain things"* (PAI_1). The fast-paced environment was felt as an attractive alternative to the usual tasks: *"it wasn't as slow as the administration usually operates, but it was fast-paced. Suddenly, you could be active. So, for the first time in a while, I really enjoyed my job again because there was real momentum"* (PAI_15).

(4) *Collaboration with stakeholders*: Crisis management is often associated more with emergency services, i.e., the operational aspect, rather than with the administrative part, which is less illustrative. On the contrary, the image of administration is pretty much the opposite of what one associates with crisis management organizations. This also affects the perception that AOS have of administrations. In Germany, this seems to have started to change only in recent years. Various prolonged situations (migration situation in 2015/2016 and again since 2022, COVID-19, energy crisis) were not classic emergency services situations but primarily required leadership through the administrative strand of crisis management.
On the one hand, administrations were very active but not immediately visible. One could rather see the classic relief organizations instead: *"Many others, of course, see on-site, (...) Red Cross, fire department, and don't even know (...) that we organized it"* (PAI_1). However, municipal administrations were key network facilitators on the local level and this corresponds with their self-perception: *"we always function as a network organization, always"* (PAI_12). Despite the health focus of the crisis, collaboration with medical professionals are not reflected as a good practice: *"the worst, (…) and I say this very critically, were the doctors. The doctors were not the pioneers. (...) They didn't want to test, they didn't want to vaccinate afterwards. They were afraid of Corona."* (PAI_15)

CONCLUSION

Drawing upon an integrated perspective on crisis management, this article emphasizes the central role of local administrations during the COVID-19 pandemic. We shed light on the significance of local administrations, which are often overlooked in both research and practice.

In the past, public administrations have been part of "forgotten tales" in crisis management research, where the focus was instead on blue light organizations and their crisis management. This is gradually being addressed as these public administrations as well as their crisis management are increasingly being included in the academic discourse. Additionally, their crucial contribution to crisis management is being recognized through crisis situations such as the 2015/16 refugee situation, the COVID-19 pandemic, or the energy crisis in Germany. The roles of local administrations are now slowly being more and more recognized and integrated into broader narratives of crisis management.

The underlying legal framework also reflects the “forgotten tales”. The pandemic has stressed the need for "new chapters and layers" within the legal paradigms to adequately capture, define and delineate crisis management. The permanent crisis mode during the pandemic provides a favorable opportunity for the adaptation and refinement of legislation, as it is currently happening in NRW. At the same time, the forgotten perspective of local governments as crisis actors was also not highlighted in the local media landscape even during COVID-19. It became apparent that local administrations were not really portrayed as crisis managers in the media and that the visible aspects or the consequences of measures tended to be described or that individual, specialized experts or important political decision-makers had their say in some articles.

In practice, the pandemic has served as a catalyst for awareness-raising and change in perception of and in local administrations. The pandemic has highlighted the shortcomings and unexplored potential of current crisis management practices and highlighted areas for improvement and innovation in public administrations. Despite this progress, experience from previous crises shows that certain lessons are not being addressed. Issues such as the follow-up and sustainability of crisis management often fall by the wayside and are overlooked after the crisis has been managed. Conversely, certain practices, such as conducting exercises, have become much more common, although this may be more of a passing trend than a permanent part of crisis preparation.

In summary, the COVID-19 pandemic has highlighted the key role of local administrations in crisis management. It has drawn attention to the legal, scientific and practical framework for crisis management and made it clear that an integrated perspective on crisis management is necessary in various crisis situations. However, not all lessons have been learned from the pandemic. With the research project and further research work, we would like to contribute to ensuring that not only immediate adjustments are made in crisis mode, but also that the sustainability of these measures is guaranteed. The experiences and adjustments made by local administrations during the pandemic provide a valuable basis for this.

ACKNOWLEDGMENTS

The research project “Local governments1 in crisis mode” is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - grant number 458544621.

REFERENCES

Aden, H. (2023). *Policing the COVID-19 Pandemic in Germany: Challenges and Learnings for Future Crises*. In M. Den Boer, E. Bervoets & L. Hak (Eds.) *Plural Policing, Security and the COVID Crisis* (pp. 139 – 160). Palgrave's Critical Policing Studies. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-031-19177-0_5

Bauer, C., & Egger, R. (2023). *Lernen aus der Pandemie: Resümee und Ausblick [Learning from the pandemic: summary and outlook.]* Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-42971-3_6

Bauer, M. W., & Grande, E. (2018). *Perspektiven der Verwaltungswissenschaft [Perspectives of Public Administration Science]*. Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783845288567

Bauer, M. W., Otto, J., & Schomaker, R. M. (2022). *„Kriseninternes Lernen" und „krisenübergreifendes Lernen" in der deutschen Kommunalverwaltung: Eine empirische Analyse von Verwaltungshandeln während der Flüchtlingskrise und der Corona-Pandemie ["Crisis-internal learning" and "cross-crisis learning" in German local government: An empirical analysis of administrative behaviour during the refugee crisis and the coronavirus pandemic]*. *Z Politikwiss, 32*, 787–804. https://doi.org/10.1007/s41358-022-00323-5

Benton, J. E. (2002). County Service Delivery: Does Government Structure Matter? *Public Administration Review*, *62*(4), 471–479. https://doi.org/10.1111/0033-3352.00200

Bogumil, J., Hafner, J., & Kuhlmann, S. (2016). Verwaltungshandeln in der Flüchtlingskrise: Die Erstaufnahmeeinrichtungen der Länder und die Zukunft des Verwaltungsvollzugssystems Asyl [Administrative action in the refugee crisis: The initial reception centres of the federal states and the future of the asylum administration system]. *Verwaltung Und Management*, *22*(3), 126–136.

Bogumil, J., & Jann, W. (2020). *Verwaltung und Verwaltungswissenschaft in Deutschland: Eine Einführung* (3rd ed.) *[Public Administration and Public Administration Science in Germany: An introduction]*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-28408-4

Eckhard, S., & Lenz, A. (2020). *Die öffentliche Wahrnehmung des Krisenmanagements in der Covid-19 Pandemie [Public perception of crisis management in the Covid-19 pandemic]* (Working Paper/Technical Report). https://kops.uni-konstanz.de/handle/123456789/49398

Ehl, F., & Wendekamm, M. (2013). Krisenmanagement als Aufgabe der politischen und administrativen Verantwortungsträger – Entscheidungen jenseits des Alltags? [Crisis management as a task for political and administrative decision-makers - decisions beyond everyday life?]. In H.-J. Lange, C. Endreß, & M. Wendekamm (Eds.), *Versicherheitlichung des Bevölkerungsschutzes* (pp. 133–146). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-02200-6_7

Hustedt, T. (2019). Verwaltung und der Umgang mit Krisen und Katastrophen [Administration and dealing with crises and disasters]. In S. Veit, C. Reichard, & G. Wewer (Eds.), *Handbuch zur Verwaltungsreform* (pp. 181–190). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-21563-7_17

Kuckartz, U., & Rädiker, S. (2022): *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* [Qualitative content analysis. Methods, Practice, Computer support] (5th ed.). Weinheim, Basel: Beltz Juventa.

Kuhlmann, S., & Franzke, J. (2022). Multi-level responses to COVID-19: crisis coordination in Germany from an intergovernmental perspective. *Local Government Studies*, *48*(2), 312–334. https://doi.org/10.1080/03003930.2021.1904398

Kuhlmann, S., & Schwab, O. (Eds.). (2017). *Starke Kommunen – wirksame Verwaltung: Fortschritte und Fallstricke der internationalen Verwaltungs- und Kommunalforschung [Strong municipalities - effective administration: progress and pitfalls in international administrative and municipal research]*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-17135-3

Lenz, A. C. (2023). *Administrative Crisis Management & Public Perceptions* [Doctoral thesis]. Universität Konstanz, Konstanz. http://nbn-resolving.de/urn:nbn:de:bsz:352-2-19nrrsazxxllo6

Möltgen-Sicking, K., & Winter, T. (2018). *Verwaltung und Verwaltungswissenschaft: Eine praxisorientierte Einführung [Administration and administrative science: A practice-orientated introduction]*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-19085-9

Neumann, T. (2020). *Handlungs(un)fähig? Kommunale Gemeindeorgane in einer Ausnahmesituation [Capable (or incapable) of acting? Municipal bodies in an exceptional situation]* (JuWissBlog No. 41). https://www.juwiss.de/41-2020/

Otto, J. (2021). *Kommunalverwaltungen in der Corona-Pandemie: Ergebnisse zweier großangelegter Befragungen in Deutschland* (WITI-Berichte No. 9) [Local authorities in the coronavirus pandemic: results of two large-scale surveys in Germany]. Deutsche Universität für Verwaltungswissenschaften Speyer. https://www.witi-innovation.de/wp-content/uploads/2021/12/2021-12-10_Otto_WITI-Bericht-9_COVID-19-Follow-up.pdf

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.), 372*, n71. https://doi.org/10.1136/bmj.n71

Pöhler, J., Bauer, M. W., Schomaker, R. M., & Ruf, V. (2020). *Kommunen und COVID-19: Ergebnisse einer Befragung von Mitarbeiter*innen deutscher Kommunalverwaltungen im April 2020* (WITI-Berichte No. 2) [*Municipalities and COVID-19: Results of a survey of employees of German local authorities in April 2020*]. urn:nbn:de:0246-opus4-66102

Sackmann, R. (2020). Sozialer Zusammenhalt bei Pandemien. Eine vergleichende Topic-Modell-Analyse zu SARS und Covid-19 [Social cohesion during pandemics. A comparative topic model analysis of SARS and Covid-19]. In N. Deitelhoff, O. Groh-Samberg, & M. Middell (Eds.), *Gesellschaftlicher Zusammenhalt: Ein interdisziplinärer Dialog* (pp. 349–375). Campus Verlag. https://doi.org/10.12907/978-3-593-44646-2

Schneider, M., & Park, K. O. (1989). Metropolitan Counties as Service Delivery Agents: The Still Forgotten Governments. *Public Administration Review*, *49*(4), 345. https://doi.org/10.2307/976844

Schönefeld, M., Schütte, P. M., Schulte, Y., & Fiedrich, F. (2023). Critical Infrastructure and Crisis Affected Actor? Investigating the Double Role of Municipal Administrations. In J. Radianti, I. Dokas, N. Lalone, & D. Khazanchi (Chairs), *ISCRAM 2023 Conference Proceedings*, Omaha, Nebraska (USA). https://idl.iscram.org/files/schoenefeld/2023/2509_Schoenefeld_etal2023.pdf

Schütte, P. M., Schulte, Y., Schönefeld, M., & Fiedrich, F. (Eds.). (2022). *Krisenmanagement am Beispiel der Flüchtlingslage 2015/2016 [Crisis management using the example of the 2015/2016 refugee situation]*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-37141-8

Schulte, Y., Schönefeld, M., Schütte, P. M., & Fiedrich, F. (2022). Herausforderungen des Krisenmanagements für öffentliche Verwaltungen [Crisis management challenges for public administrations]. In A. H. Karsten (Ed.), *Resilienz und Pandemie* (1st ed., pp. 130–137). W. Kohlhammer GmbH.

Thomas, J., O'Mara-Eves, A., Harden, A., & Newman, M. (2017). Synthesis Methods for Combining and configuring textual or mixed methods data. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 181–210).

Ziekow, J. (Ed.). (2018). Verwaltungspraxis und Verwaltungswissenschaft. *Schriften der Deutschen Sektion des internationalen Instituts für Verwaltungswissenschaften: Band 41 [Administrative practice and administrative science]*. Nomos. https://doi.org/10.5771/9783845290522

Uncovering Impact of Innovation: Continuous Stakeholder Engagement through Scenario-based Systems Engineering

Jens Pottebaum
Paderborn University
Heinz Nixdorf Institute
Chair for Product Creation
jens.pottebaum@hni.upb.de

Marcel Ebel
Paderborn University
Heinz Nixdorf Institute
Chair for Product Creation
marcel.ebel@hni.upb.de

Iris Gräßler
Paderborn University
Heinz Nixdorf Institute
Chair for Product Creation
iris.graessler@hni.upb.de

ABSTRACT

The importance of extreme weather situations is increasing due to their number and, above all, their impact on stakeholders in emergency response. They are characterized by cascading effects with global and local interdependencies. Extreme data must be included as a basis for decision-making. The impact in emergency response depends on diverse, multidisciplinary competencies required to interpret information. Scenarios are used in various forms of preparation: in exercises, but also for the design of information systems and validation. Based on literature, this article brings together different types of scenarios and related work in the field of Model-Based Systems Engineering. Using an exemplary case relating to possible pluvial urban floods, the added value resulting from a focus on the impact of innovative solutions is discussed. It is shown that the use of scenarios helps to make the desired impact assessable for decision-makers in all phases of research and development projects.

Keywords

systems, scenarios, engineering, models, innovation

INTRODUCTION

Extreme weather situations are increasing due to climate change (Lee et al., 2023). Weather-induced emergencies are characterized by cascading effects with global and local interdependencies (Schauwecker et al., 2019). In hazard analysis and response, it is therefore necessary to deal with extreme data that is obtained from heterogeneous data sources and evaluated using various methods, including machine learning (cf. (Abily et al., 2020)). Examples include global data such as geodata, traffic information, satellite images and weather forecasts on the one hand, and local data from sensor systems such as weather stations, ground-based or flying robots (UAV/UGV) on the other (Kruijff-Korbayova et al., 2021). Extreme weather therefore places extreme demands on information and communication technologies (ICT) (Terti et al., 2019) and self-protection of the population (Gräßler et al., 2018). Effective support for all stakeholders requires the provision of weather forecasts and, above all, constantly updated impact forecasts as the emergency progresses, in order to prepare for such an emergency and its development over time (e.g., (Sutanto et al., 2019)). This becomes even more important the less specialist expertise is available in a command and control post in areas such as meteorology, geology or sensor technology.

WiP Paper – Impact-driven Emergency and Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

The management of critical situations in real time with extreme and complex data must therefore be based on an assessment of the information quality of extreme data. Different fields of view and sampling frequencies of sensor technology, latencies in communication, gaps in data collection, differences in data formats, etc. characterize "extreme data" - the term "big data" is therefore avoided due to the too narrow reference to volume. In the decision-making process, an awareness must be created that uncertainty is always inherent here. Indicators of data and information quality offer a possible approach (Eppler, 2006). The achievable effect depends on the application context in different command and assistance centers as well as the available infrastructure with devices for visualization, interfaces of weather services, sensor systems and rescue robotics. This article is therefore based on the assumption that the consistent use of scenarios is crucial for the successful impact of innovative solutions in the application. This can be based on the approaches of Model-Based Systems Engineering (MBSE) (Gräßler et al., 2023). While requirements can be represented in previous system models, there is no explication of narrative knowledge. In scenarios, these can be discussed and referenced (cf. (Friberg et al., 2010)).

The article summarizes different types of scenarios based on the literature in the following chapter. Related work in the field of MBSE is adopted for application in emergency management. The concept for the conceptualization and abstraction of different types of scenarios covers the entire chain from domain exploration to the evaluation of the impact of a technology in this domain. Using a case study relating to possible pluvial urban flooding situations, the added value that arises along the chain from actual practice scenarios to target application scenarios and validation scenarios is discussed. In the conclusion, it is shown that the use of scenarios can help to make the desired effect of innovations explicit early on in development and research projects and thus make them assessable for decision-makers.

STATE-OF-THE-ART IN SCENARIO REPRESENTATION

The aim of scenarios is to transform statements into descriptive, intuitively usable auxiliary artifacts. They thus combine the illustrative perspective with a purpose-related, constructive view. The challenge is the qualitative description based on subjective descriptions and perceptions. These can be related to the problem and solution space. With regard to the problem space in particular, it becomes clear that the need for information extends well beyond the domain of product development. Accordingly, information from different domains is integrated here. Distortions as a result of the scenario form must be taken into account (Sabatucci et al., 2015).

Scenarios in Research on ICT in Impact-Oriented Emergency Management

Systematically researching the ISCRAM library following the principles of the PRISMA statement (Page et al., 2021), several notions of “scenarios” become obvious. This is supported by cross-checking with Web of Science and Scopus databases. Amongst others, there is a variety of work that applies scenarios in (virtual) simulations resp. exercises. Further references are available that make use of scenarios to model and understand a certain domain of interest or type of hazard, like people evacuation, food supply or humanitarian assistance. When searching for an exact match of scenarios and systems engineering or related terms, there are not direct matches. Extending the scope, some papers can be identified that support the relevance of scenarios along design and engineering activities for ICT systems in emergency and crisis management. Intending an advanced User Experience, TOULOUM et al. propose to incorporate service touchpoints explicitly in scenarios (Touloum et al., 2013). Based on structured scenarios from overall scenarios to activity diagrams and interaction models (cf. (Friberg et al., 2010)), touchpoints are extracted representing those system elements that require UX design interventions. GARCIA-ARISTIZABAL et al. introduce their work on Cascading Effects Scenarios with several case studies (Garcia-Aristizabal et al., 2015). Their work combines analytical elements when describing past emergencies through scenarios (as-is scenarios) with prescriptive use cases when targeting supportive measures like ICT systems. RAMIREZ et al. focus on risk analysis based on scenarios (Ramirez de la Huerga, Miguel et al., 2015). Even though this is not directly linked to Systems Engineering, the presented approach can help to prioritize scenarios targeting risks of cascading events. This could be coupled with Systems Engineering approaches targeting similar risk-oriented objectives (Gräßler et al., 2022). On a different layer, POLIPARKUS et al. research on collaboration in (virtual) scenario generation (Poliparkus et al., 2018). Understanding Systems Engineering as a set of tasks that needs to be fulfilled in a collaborative manner, especially the findings regarding feedback loops and finetuning of scenarios to ensure realistic settings are considered in the paper at hand.

Criteria to Categorize Scenario Types

Process-related scenarios usually include events and/or activities in narrative form. One example is the description of an actual process within which problems are to be identified. Situation-related scenarios, on the other hand, describe states as intended or anticipated situations. One example is future scenarios that anticipate possible

futures. They can also include a temporal dimension that describes the achievement of this situation or a further development. Selection scenarios are used to support decision-making and therefore have a reference to a decision-making situation. In the application perspective, the use of a product is assumed, which must be available accordingly. ROLLAND et al. recommend a classification of scenarios based on form, content, purpose and life cycle (Rolland et al., 1998, p. 45; Weidenhaupt et al., 1998, p. 35). FILIPPIDOU identifies form- and content-related comparison criteria for the classification of scenarios based on a literature analysis (Filippidou, 1998, p. 5) (extended based on (Jarke et al., 1998; Pohl, 2010, p. 148; Rolland et al., 1998; Weidenhaupt et al., 1998, p. 37)).

Table 1. Criteria to categorize scenario types

Layer 1	Layer 2	Values
form	degree of formalization	informal to formal
	form of presentation	statistical to animation/simulation
	stakeholder input	purely synthetic to intense (workshops etc.)
	abstraction hierarchy	existing (e.g., use case models) to not existing (unstructured, without underlying schema)
	complexity hierarchy	Existing or non-existing classification with regard to a combination of generalization and specialization
	method of generation	automated (e.g., (Bagschik et al., 2018, p. 1816)), reference models, …
	method of analysis	observatory, argumentative, automated
	intention	descriptive, explorative or explanatory (Pohl, 2010, p. 148)
content	concreteness	„3ague“ (e.g., future scenarios) to detailed (e.g., test scenarios)
	boundedness	specific situation to “big picture” (overarching, generic view), system-related differentiation into system-internal, interaction and environment/context scenarios (Jarke et al., 1998; Pohl, 2010, p. 148)
	plausibility	differentiation between desired/positive scenarios and scenarios to be avoided/negative scenarios or scenarios that contribute to or jeopardize objectives (incl. misuse scenarios) (Pohl, 2010, p. 148)
	visualization	current/actual/indicative vs. planned/optional scenarios (cf. (Jarke et al., 1998, p. 157)), also differentiation between instance and type scenarios (Pohl, 2010, p. 148); supplements are anticipated scenarios
life cycle	criticism	not done – testable/possible – tested/iterated
	evolution	temporary scenarios with dynamics vs. persistent scenarios
	alternatives	alternatives: main, alternative and exceptional scenarios (Pohl, 2010, p. 148)

Requirements Engineering: Practice und Interaction Scenarios

Requirements engineering is a cooperative, iterative and incremental process (Pohl, 2010). It includes eliciting, analyzing and structuring, documenting, negotiating and specifying as well as managing and validating requirements (Pohl, 2010, 41ff; Rupp, 2014). The use of scenarios as auxiliary artifacts is taken up in Scenario-Based Requirements Engineering (A. G. Sutcliffe et al., 1998). Scenario types in the development of e.g. requirements and prototypes require a concrete, possibly simulatable and at least semi-formal form of representation. They are used a) to represent a current state or a current situation or b) to describe a desired state or a desired situation (Pohl, 2010, 149ff). ROSSON and CARROLL emphasize problem scenarios in the first case (Rosson & Carroll, 2008, p. 26). In the future-oriented case, the focus can be on interaction with the technical system to be developed (interaction scenario, e.g. in the form of user stories or use cases). Anticipated errors in product use (misuse scenario), deliberately incorrect use (abuse scenario) and even threat scenarios can be modeled here (see, e.g., (Gräßler et al., 2020; Tuma et al., 2018)).

RELATED WORK IN MODEL-BASED SYSTEMS ENGINEERING

The largest group of related work deals with metadata modeling for scenarios in requirements development. ARRUDA et al. develop the BD-REAM in the form of an entity-relationship model (Arruda, 2020; Arruda & Madhavji, 2017). Big data requirements are placed in the context of software, system and quality requirements, test and business cases. In a software implementation (QualiBD), for instance the completeness of a requirements set is checked (Arruda, 2020, p. 157). BERKOVICH et al. develop a domain mapping matrix (DMM) for mapping relations of different requirement types (Berkovich et al., 2014). Even though "needs" are explicitly addressed, a formal mapping seems to start later at the level of customer requirements. The Scenario-based Requirement Model (SBRM) from Xi et al. models scenarios from the perspective of requirements: A requirement is visualized by scenarios, processes and rules contained therein (Xi et al., 2019). This perspective is deepened by the ontology development in this domain (see, e.g., literature studies (Castañeda et al., 2010; Dermeval et al., 2016)). ReqIF, as mentioned above, contains a data schema for "SpecObjects" (requirements) and "Specification" (requirement set) (Object Management Group [OMG], 2016). There is no abstraction to the semantic level here, only recommendations for the exchange of requirements artifacts. TUDORACHE models dependencies between development artifacts including requirements. It uses the semantic level to check consistency, conformity and compliance with the specification by means of "design models" (Tudorache, 2006). The SoftWiki Ontology for Requirements Engineering (SWORE) is a central work in terms of metadata modeling (Riechert & Berger, 2009; Riechert et al., 2007). SIEGEMUND is more specific in the area of the representation of metrics (also in the form of qualitative SoftMetric), test cases and obstacles (Siegemund, 2014, p. 60). He maps the response to uncertainty with reference to risks, without modeling the uncertainty directly. Individual works focus on requirements change management. ALSANAD et al. present an ontology that declares the requirements, parts of requirements and requirements sets for the Aspect_of_Change as reference objects for changes (Alsanad et al., 2019, p. 49357). Uncertainty is part of the motivation in some works (e.g. (Heling et al., 2018)), but only in a few approaches explicitly part of the solution approach. FEHLMANN et al. do not focus on the representation of uncertainties, but describe the propagation of "customer needs" to use cases or test cases (measurable e.g. via the number of bugs) (Fehlmann & Santillo, 2007, p. 204). The solution approaches mentioned are presented in Figure 1 in a comparative overall context. Work that focuses on the work processes for creating these scenario artifacts without explicit modeling is not considered here (see, e.g., (Georg et al., 2015)).

● focus area ◑ partial approach ○ only initial - no contribution **Criterion**	Formalized management of scenario data	Traceable RE metadata model	Traceable interfaces to other development artifacts (internal)	Traceable interfaces to underlying sources (external)	Explicit representation of uncertainty
Arruda et al. (BD-REAM, 2017/2020)	○	●	●	-	-
Berkovich et al. (2014)	-	◑	●	-	-
Xi et al. (SBRM, 2019)	●	○	-	-	-
Kügler et al. (2018)	-	◑	◑	●	-
ReqIF (2016)	-	●	○	-	-
Tudorache (2006)	-	◑	●	-	-
Riechert et al. (SWORE, 2007/2008)	◑	●	○	-	-
Siegemund (2014)	◑	●	-	-	○
Alsanad et al. (2019)	-	○	●	-	-
Fehlmann et al. (2007)	-	○	○	○	●
Binger (2006)	●	-	-	◑	◑

Figure 1. Related work on impact-oriented, scenario-based Systems Engineering

CLASSIFICATION AND ABSTRACTION OF SCENARIOS

Along Systems Engineering phases, different scenario types are relevant (Anggreeni & Voort, 2007), which are incorporated in Figure 2 based on criteria in Table 1: Illustrative scenarios in the form of explorative scenarios (e.g. (potential) stakeholder stories) and actual or intended practical scenarios serve to reduce ignorance in early phases of the V model. Interaction, detailed scenarios and validation/test scenarios are used in later phases, also closing feedback loops across engineering iterations.

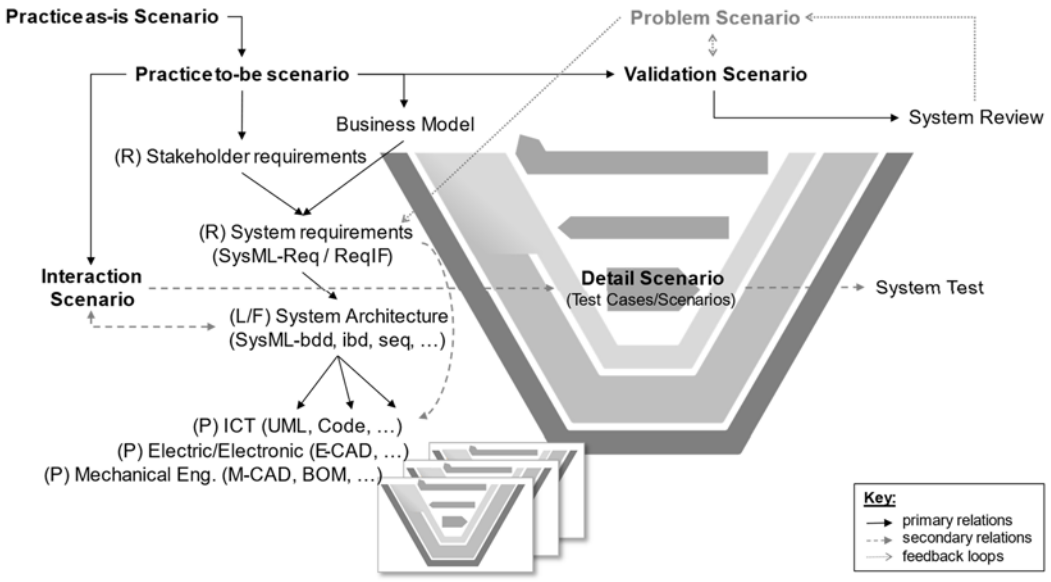


Figure 2. Utilization of scenarios along (Model-Based) Systems Engineering tasks

Classification of Scenarios

Based on the classification scenarios, basic assumptions can be made with regard to the dependence on information quality (Figure 3). Modeling and representation in the form of information resources is usually carried out using standard software as an authoring tool (e.g. Microsoft Office). Special tools can be used in scenario management, but are rarely in productive use (see as an example (Anggreeni & Voort, 2009)).

● high
◑ dependent on implementation
○ low

			Possible degree of formalization	Dynamics in the form of presentation (simulation)	Independence from stakeholder contribution	Existing abstraction hierarchy	Specificity in complexity hierarchy	Automated creation	Automated analysis	Intention (observational to explanatory)	Concreteness	Boundedness	Degree of plausibility assurance	Degree of visualization	Degree of criticism received	Dynamics of a possible evolution
1	Future Scenario		○	○	○	○	○	○	○	●	○	○	◑	○	◑	●
2	Explorative Scenario		○	○	○	○	◑	◑	◑	◑	○	◑	○	●	○	◑
3	Practice Scenario	as-is	◑	◑	○	◑	◑	◑	◑	○	◑	◑	◑	●	◑	○
4		to-be	◑	◑	◑	◑	◑	◑	◑	◑	◑	◑	●	◑	◑	◑
5	Interaction Scenario		●	◑	◑	●	◑	◑	◑	◑	◑	◑	●	◑	◑	◑
6	Problem Scenario		●	●	◑	●	●	◑	◑	○	◑	●	●	●	◑	○
7	Detail Scenario		●	●	◑	●	●	◑	◑	●	●	●	●	●	◑	◑
8	Validation Scenario		●	●	◑	●	●	●	◑	●	●	●	●	●	●	○

Figure 3. Classification of scenarios with corresponding interrelations

Abstraction of Scenarios

The definition of scenarios is initially based on the definition of a scenario type and an associated scenario form. Figure 4 shows a corresponding procedure with reference to data sources. These include free text, structured text, images, animations or simulations, state charts, process models, activity and interaction diagrams (Weidenhaupt et al., 1998, p. 37). For these types of scenarios, formalized and standardized forms are defined to varying degrees. Depending on the type of scenario, priorities are set: For example, while state machines (standardized as part of UML and SysML, among others) emphasize the relationship between events and states and activities are secondary, the latter are the focus of process models (e.g. UML/SysML, eEPK and BPMN) with only implicitly recognizable state changes. In line with this focus, data and information are collected and processed in order to design scenarios. These are purpose-related. Validation is recommended as a prerequisite for reuse (A. G. Sutcliffe et al., 1998, p. 1074). One example is the use for requirements elicitation and later for the derivation of test cases.

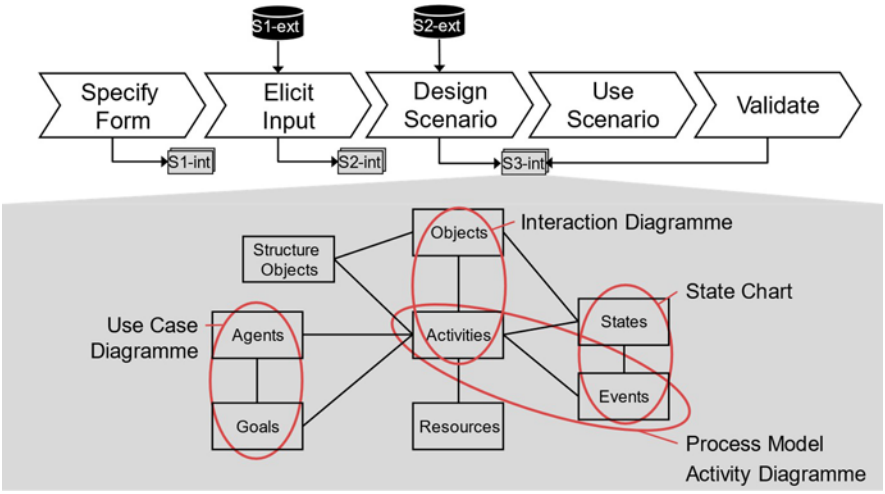


Figure 4. Scenario definition, elements and types (based on (A. Sutcliffe, 1998, p. 51; A. G. Sutcliffe et al., 1998, p. 1074))

The input (S1-ext) is collected, for example, through observation or questioning, content analysis on the basis of specifications including underlying regulations or the interpretation of experiences from previous product generations. As a result of this step (S2-int), ignorance with regard to the initial data must be taken into account as well as ignorance with regard to the relevance for the product to be developed. When designing scenarios, (partial) scenarios that have been defined for comparable products and product types can be reused (S2-ext). The indeterminacy in the design of scenarios as an active, decision-making activity results from the selected form (S1-int) and the content depicted therein (S3-int). Both forms of uncertainty must be examined in the validation.

Conceptualization of Practice and Interaction Scenarios

A metadata model to support scenario-based requirements engineering must support the use of scenarios in the activities from elicitation to analysis and validation of requirements. Accordingly, the differentiated integration of scenarios in requirements engineering must be realized on the one hand, and integration into the higher-level system model on the other. In addition, the projections of future requirements and the risk of change to requirements are challenges in requirements analysis (Graessler et al., 2020).

Requirements should be modeled in a (semi-)structured way. Templates prepare the formalization of metadata and can be used as a basis for conceptualization. Natural language is used for the representation of requirement descriptions. Requirement attributes are defined in the SysML requirement artifact specification, among others. (INCOSE, 2014). Templates are available in the form of the Volere template for requirements (Robertson & Robertson, 2013) and the SOPHIST schemas for requirement descriptions (Rupp, 2014). Its use ensures both a high degree of structuring and flexibility for intuitive descriptions. The semantic model is extended by an ontology for requirements engineering, which is based on the „SWORE – SoftWiki Ontology for Requirements Engineering“ (Dermeval et al., 2016; Riechert & Berger, 2009; Riechert et al., 2007). According to ReqIF (OMG, 2016, p. 12) and the concept introduced by BEVAN et al. (Bevan et al., 2018, p. 8), a distinction between stakeholder requirements and system requirements is recommended. Requirements are logically bundled into sets, which in turn are documented in specifications at defined points in time or events.

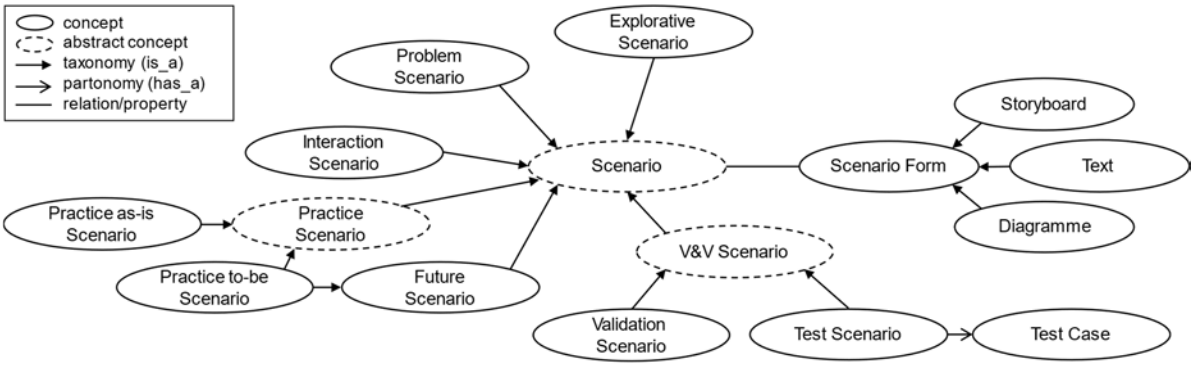


Figure 5. Representation of scenario types and their interrelations

The basis of the conceptual model presented in Figure 5 is attributed by the above-mentioned properties. As special scenario form, practice scenarios are derived as indicative actual or optative target scenarios. Future scenarios depict needs or serve to illustrate business cases. They have an explanatory effect or are used for exploratory analysis (Anggreeni & Voort, 2007). This prepares the ground for both market-oriented and order-related use. Interaction scenarios are derived in a solution-oriented manner, which can also be expressed in the negative form of problem or even misuse scenarios (Pohl, 2010, p. 148). Validation scenarios in the broader sense mentioned above are concretized in terms of Verification and Validation (V&V). Such V&V scenarios are understood as an abstract concept which is detailed by validation scenarios (involving end users or representatives) and test scenarios. Test scenarios are either created based on test cases (e.g., when field trials are scheduled from scratch), or test cases are incorporated into existing test scenarios (e.g., when an exercise is already scheduled).

ADOPTION FOR INNOVATION IN WEATHER-INDUCED EMERGENCIES

The value of traceable scenario utilization of discussed based on two successive exemplary projects. They focus on the algorithms and technical capabilities indicated by the layers in Figure 5. In general, data needs to be integrated including both the current situation (data collected until t_0) and a projected situation (anticipating situation at t_{+m}). Challenges subsume spatial and temporal domains. All layers are based on efficient data acquisition. They include the entire chain from processing close to the sensor, extraction of information and its visualization. It is crucial to ensure explainability in the user interface. Instead of technical explainability, which is understood as technical measures to explain the underlying model, solutions are required to provide decision-makers with information that they perceive as "sufficiently explained". In this case example, rescue robotics (Kruijff-Korbayova et al., 2021) are combined with a Geo Information System (GIS) which was built on the basis of results from the ANYWHERE project (see Figure 6, cf. (Abily et al., 2020; Van Lanen et al., 2019)). Users include civil protection, but also public or private institutions such as water management or the management of flood evacuation plans in campsites. This functional basis is available to implement and discuss the chain of interdependent scenarios. Figure 6 shows the processing levels from data at the bottom to visualized information on top of the stack. The vision of CREXDATA is to develop a generic platform for real-time critical situation management including flexible action planning and agile decision making over data of extreme scale and complexity. Temperatures, videos, forest fire index and wind conditions are indicated as examples of data at the bottom. All of these data types are processed using methods such as trained classifiers or Complex Event Forecasting (cf. (Pottebaum et al., 2012)). An 'explainability layer' in the sense of eXplainable AI (XAI) complements content with additional information that is understandable to humans. This creates the basis for applying visualization approaches.

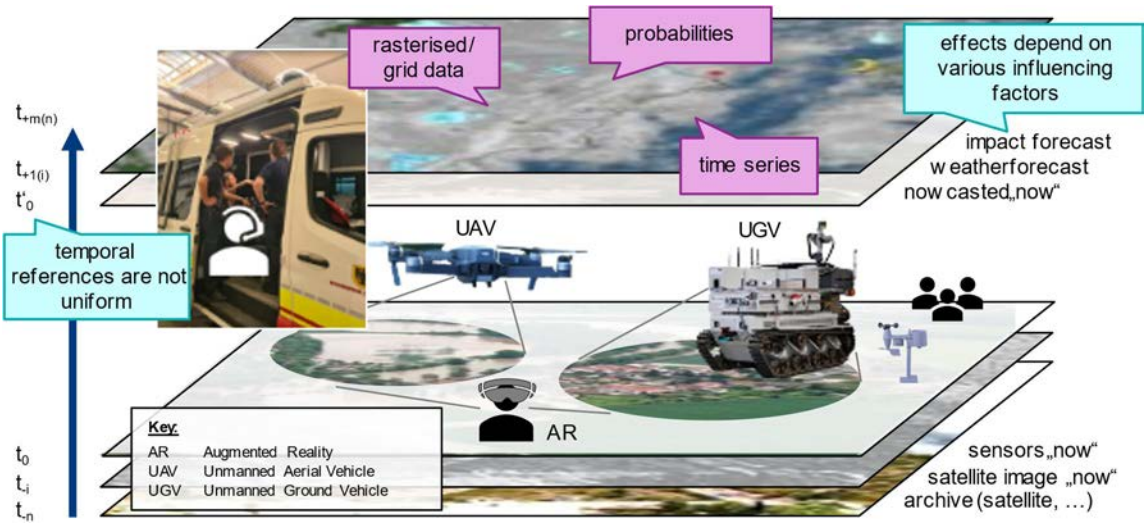


Figure 6. Exemplary case of an ICT solution to support multi-level decision-making in extreme weather scenarios considering global data, for example provided by UAV/UGVs and meteorological services (Pottebaum et al., 2023).

The design of the system is extremely challenging due to several reasons. In addition to the general challenge that there is a large gap between technical opportunities and technologies actually deployed, such reasons include:

- complexity of stakeholder landscape and use cases across functions/institutions (breadth of use cases) and along hierarchy of command (depth of use cases)
- sparsity of extreme weather events, implying that only few experienced stakeholders can be acquired
- variety of domains, often requiring deep knowledge up to scientific competences (like meteorology)
- intellectual Property across data sources, with heterogenous users and their licenses
- variety of technical interfaces
- uncertainty of data along the data processing pipeline

Therefore, the concept of scenarios was applied from early explorative phases through requirements engineering to planning of evaluation with laboratory settings and field trials. Figure 7 presents an overview of the approach incorporating different types and forms of scenarios utilized in impact-oriented research.

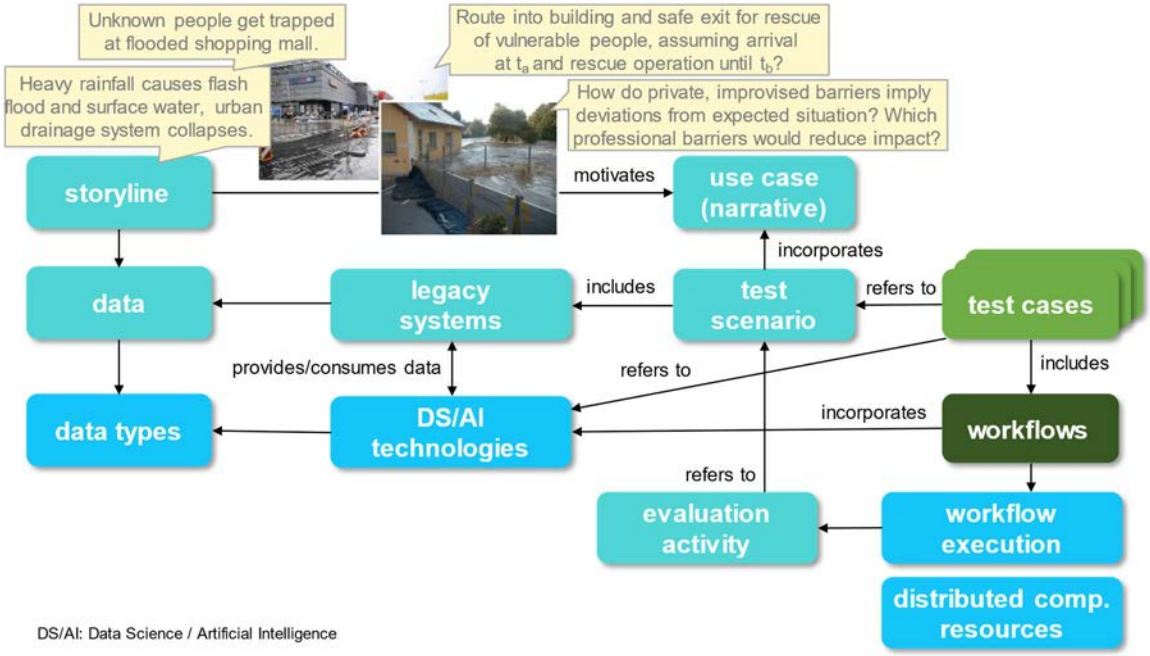


Figure 7. Adoption of various types of scenarios in research on utilizing extreme data with high uncertainties for precise decision support in weather-induced emergencies

The explorative phase of the CREXDATA project was supported by informal "storylines", established in the form of storyboards. Two storylines are created in intense stakeholder meetings to detail the general "pluvial urban flooding" scenarios. In the first storyline, the evacuation of a shopping mall in Dortmund (Germany) with heterogenous visitors is required. In the second storyline, water management with barriers is needed to assess and to avoid threats to traffic and buildings in Innsbruck (Austria). Individual steps within these storyboards become referrable by linking them to simple activity diagrams. Process modelling with Business Process Model and Notation (BPMN) simplifies uptake of existing process models wherever possible. From such storylines, an early exploration is possible regarding relevant data (cf. distinction of global and local data introduced above) and corresponding data types. This is essential in projects that envisage uptake of Data Science (DS) and Artificial Intelligence (AI) technologies. For instance, for technologies like Complex Event Forecasting (CEF), time series data is required in combination with logical modelling of relevant event types to be recognized resp. forecasted. The initial indication on potential data sources enables a two-sided approach: Potential event types can be prescribed by users parsing storylines stepwise, and data sets can be analyzed by CEF partners to identify patterns in data without in-depth domain knowledge. Problem-oriented storylines as explorative scenarios and practice as-is scenarios are transformed into practice to-be scenarios and use case narratives. Such narratives still subsume sketches and natural language bodies but described in a solution-oriented way. These artefacts incorporate knowledge about the technologies to be integrated into the system under development. Bridging system design and system Verification & Validation (V&V), test scenarios are derived. Each test scenario subsumes specific test cases. In research on DS/AI data processing, test cases refer to specific data processing workflows designed using Altair RapidMiner Studio. Test scenarios are utilized with respect to evaluation activities in a two-fold approach: On the one hand, scheduled exercises of stakeholders are identified and checked for applicability in the CREXDATA project. Often, exercises conducted by stakeholders like civil protection authorities, technical relief agencies or fire departments anyhow are most realistic measures for gathering feedback. Researchers need to map exercise objectives and evaluation objectives. On the other hand, field trials are scheduled with the dedicated purpose of evaluation. Practicing procedures is a side-effect for stakeholders then. For the CREXDATA project itself, the mapping of evaluation activities to workflows through test scenarios and test cases ensures traceability of results, enables actual conclusions from application to data processing foundations, and simplifies communication among all partners.

CONCLUSIONS

The article summarizes different types of scenarios based on the literature. A focus is set to scenario types utilized in Requirements Engineering, as there is a variety of background available to gather feedback from stakeholders. Related work in the field of MBSE is adopted for application in emergency management. Scenario types subsume problem-oriented types like exploratory and practice as-is scenarios, as well as solution- and, as such, impact-oriented types like practice to-be and interaction scenarios. The concept for the conceptualization and abstraction of different types of scenarios covers the entire chain from domain exploration to the evaluation of the impact of a technology in this domain. It is based on the V model of Systems Engineering. Using a case study relating to possible pluvial urban flooding situations, the added value that arises along the chain from actual practice scenarios to target application scenarios and validation scenarios is discussed based on two successive exemplary projects. As a conclusion, the mapping along the entire chain from explorative storylines through practice scenarios to evaluation activities ensures traceability of results. The linkage of these scenarios to workflows through test scenarios and test cases enables actual conclusions from application to data processing foundations and simplifies communication among all partners.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union's Horizon Europe Program under the CREXDATA Project, grant agreement n° 101092749.

REFERENCES

Abily, M., Gourbesville, P., Carvalho Filho, E. de, Llort, X., Rebora, N., Sanchez, A., & Sempere-Torres, D. (2020). Anywhere: Enhancing Emergency Management and Response to Extreme Weather and Climate Events. In P. Gourbesville & G. Caignaert (Eds.), *Springer Water. Advances in Hydroinformatics* (pp. 29–37). Springer Singapore. https://doi.org/10.1007/978-981-15-5436-0_3

Alsanad, A. A., Chikh, A., & Mirza, A. (2019). A Domain Ontology for Software Requirements Change Management in Global Software Development Environment. *IEEE Access*, *7*, 49352–49361. https://doi.org/10.1109/ACCESS.2019.2909839

Anggreeni, I., & Voort, M. C. van der. (2007). *Tracing the Scenarios in Scenario-Based Product Design: a study to support scenario generation* (TR-CTI). University of Twente.

Anggreeni, I., & Voort, M. C. van der. (2009). Supporting Scenario-Based Product Design and Its Adapters: An Informal Framework for Scenario Creation and Use. In M. J. Smith & G. Salvendy (Eds.), *Lecture Notes in Computer Science: Vol. 5617. Symposium on Human Interface 2009, held as part of HCI International 2009, San Diego, CA, USA, July 19 - 24, 2009 ; proceedings, part I* (Vol. 5617, pp. 217–226). Springer. https://doi.org/10.1007/978-3-642-02556-3_25

Arruda, D. (2020). *Requirements Engineering in the Context of Big Data Software Applications Application* [Dissertation]. The University of Western Ontario, Ontario. https://ir.lib.uwo.ca/etd/6808

Arruda, D., & Madhavji, N. H. (2017). Towards a requirements engineering artefact model in the context of big data software development projects: Research in progress. In J.-Y. Nie, Z. Obradovic, T. Suzumura, R. Ghosh, R. Nambiar, & C. Wang (Eds.), *2017 IEEE International Conference on Big Data* (pp. 2314–2319). IEEE. https://doi.org/10.1109/BigData.2017.8258185

Bagschik, G., Menzel, T., & Maurer, M. (2018). Ontology based Scene Creation for theDevelopment of Automated Vehicles. In IEEE Intelligent Transportation Systems Society (Ed.), *2018 IEEE Intelligent Vehicles Symposium* (pp. 1813–1820). IEEE.

Berkovich, M., Leimeister, J. M., Hoffmann, A., & Krcmar, H. (2014). A requirements data model for product service systems. *Requirements Engineering*, *19*(2), 161–186. https://doi.org/10.1007/s00766-012-0164-1

Bevan, N., Carter, J., Earthy, J., Geis, T., & Harker, S. (2018). What Are User Requirements? Developing an ISO Standard. In M. Kurosu (Ed.), *Lecture Notes in Computer Science. Human-Computer Interaction. Theories, Methods, and Human Issues* (Vol. 10901, pp. 3–13). Springer International. https://doi.org/10.1007/978-3-319-91238-7_1

Böhm, F., Fähling, J., Huber, M. J., Leimeister, J. M., & Krcmar, H. (2011). Scenario planning for innovation development: an overview of different innovation domains. In *R&D Management Conference.* Symposium conducted at the meeting of Linköping University, Norrköping.

Castañeda, V., Ballejos, L., Caliusco, L., & Galli, R. (2010). The Use of Ontologies in Requirements Engineering. *Global Journal of Research in Engineering*, *10*(6), 2–8.

Dermeval, D., Vilela, J., Bittencourt, I. I., Castro, J., Isotani, S., Brito, P., & Silva, A. (2016). Applications of ontologies in requirements engineering: A systematic review of the literature. *Requirements Engineering*, *21*(4), 405–437. https://doi.org/10.1007/s00766-015-0222-6

Eppler, M. J. (2006). *Managing Information Quality: Increasing the Value of Information in Knowledge-intensive Products and Processes* (Second Edition). Springer. http://dx.doi.org/10.1007/3-540-32225-6

Fehlmann, T., & Santillo, L. (2007). Uncertainty of Software Requirements. In T. Dekkers (Chair), *Software Measurement European Forum*, Centro Congressi FRENTANI, ROME (ITALY).

Filippidou, D. (1998). Designing with scenarios: A critical review of current research and practice. *Requirements Engineering*, *3*(1), 1–22. https://doi.org/10.1007/BF02802918

Friberg, T., Birkhäuser, B., Pottebaum, J., & Koch, R. (2010). Using scenarios for the identification of real-world events in an event-based system. In Chris Zobel & B. T. S. French (Chairs), *International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, Seattle.

Garcia-Aristizabal, A., Polese, M., Zuccaro, G., Almeida, M., & Aubrecht, C. (2015). Improving emergency preparedness with simulation of cascading events scenarios. In L. Palen, M. Büscher, T. Comes, & A. L. Hughes (Eds.), *12th Proceedings of the International Conference on Information Systems for Crisis Response and Management, Krystiansand, Norway, May 24-27, 2015*. ISCRAM Association.

Georg, G., Mussbacher, G., Amyot, D., Petriu, D., Troup, L., Lozano-Fuentes, S., & France, R. (2015). Synergy between Activity Theory and goal/scenario modeling for requirements elicitation, analysis, and evolution.

Information and Software Technology, *59*, 109–135. https://doi.org/10.1016/j.infsof.2014.11.003

Graessler, I., Oleff, C., & Scholle, P. (2020). Method for Systematic Assessment of Requirement Change Risk in Industrial Practice. *Applied Sciences*, *10*(23), 8697–8725. https://doi.org/10.3390/app10238697

Gräßler, I., Bodden, E., Pottebaum, J., Geismann, J., & Roesmann, D. (2020). Security-Oriented Fault-Tolerance in Systems Engineering: A Conceptual Threat Modelling Approach for Cyber-Physical Production Systems. In *Advances in Intelligent Systems and Computing. Advanced, Contemporary Control* (Vol. 1196, pp. 1458–1469). Springer International. https://doi.org/10.1007/978-3-030-50936-1_121

Gräßler, I., Oleff, C., & Preuß, D. (2022). Proactive Management of Requirement Changes in the Development of Complex Technical Systems. *Applied Sciences*, *12*(4), 1874. https://doi.org/10.3390/app12041874

Gräßler, I., Pottebaum, J., & Scholle, P. (2018). Influence Factors for Innovation in Digital Self-Preparedness Services and Tools. *International Journal of Information Systems for Crisis Response and Management*, *10*(1), 20–37. https://doi.org/10.4018/IJISCRAM.2018010102

Gräßler, I., Wiechel, D., Koch, A.-S., Sturm, T., & Markfelder, T. (2023). Methodology for Certification-Compliant Effect-Chain Modeling. *Systems*, *11*(3), 154. https://doi.org/10.3390/systems11030154

Heling, B., Schleich, B., & Wartzack, S. (2018). Robust-Design-Optimization of mechanisms based on kinematic requirements considering uncertainties. *Procedia CIRP*, *75*, 27–32. https://doi.org/10.1016/j.procir.2018.04.048

INCOSE. (2014). *Systems Engineering Handbook: A Guide for System Life Cycle Processes And Activities.*

International Institute of Business Analysis (Ed.). (2009). *A guide to the business analysis body of knowledge: (BABOK® guide)* (Version 2.0). IIBA.

Jarke, M., Bui, X. T., & Carroll, J. M [John M.] (1998). Scenario Management: An Interdisciplinary Approach. *Requirements Engineering*, *3*(3-4), 155–173. https://doi.org/10.1007/s007660050002

Kruijff-Korbayova, I., Grafe, R., Heidemann, N., Berrang, A., Hussung, C., Willms, C., Fettke, P., Beul, M., Quenzel, J., Schleich, D., Behnke, S., Tiemann, J., Guldenring, J., Patchou, M., Arendt, C., Wietfeld, C., Daun, K., Schnaubelt, M., Stryk, O. von, Surmann, H. (2021). German Rescue Robotics Center (DRZ): A Holistic Approach for Robotic Systems Assisting in Emergency Response. In *2021 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)* (pp. 138–145). IEEE. https://doi.org/10.1109/SSRR53300.2021.9597869

Lee, H., Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P. W., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W. W., Connors, S., Denton, F., Diongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., Péan, C. (2023). *IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.* https://doi.org/10.59327/IPCC/AR6-9789291691647

Object Management Group (OMG) (2016). *Requirements Interchange Format (ReqIF):* Version 1.2 (formal/2016-07-01). http://www.omg.org/spec/ReqIF/1.2

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*(1), 89. https://doi.org/10.1186/s13643-021-01626-4

Pohl, K. (2010). *Requirements engineering: Fundamentals, principles, and techniques*. Springer.

Poliparkus, S., Ley, T., & Poom-Valickis, K. (2018). Collaborative Authoring of Virtual Simulation Scenarios for Assessing Situational Awareness. In Anouck Adrot, Rob Grace, Kathleen Moore, & Christopher Zobel (Chairs), *ISCRAM*, Blacksburg, VA, USA.

Pottebaum, J., Artikis, A., Marterer, R., & Paliouras, G. (2012). User-Oriented Evaluation of Event-Based Decision Spport Systems. In *IEEE 24th International Conference on Tools with Artificial Intelligence* (pp. 162–169). IEEE Computer Society. https://doi.org/10.1109/ICTAI.2012.30

Pottebaum, J., Rechberger, C., Hieb, M., Gräßler, I., & Resch, C. (2023). Extremwettersituationen in alpinen Gebieten: Management kritischer Situationen in Echtzeit bei extremen und komplexen Daten. In Disaster Competence Network Austria (Chair), *Tagungsband der Fachtagung Katastrophenforschung 2023*.

Ramirez de la Huerga, Miguel, Banuls, V. A., & Turroff, M. (2015). A Scenario-based approach for analyzing complex cascading effects in Operational Risk Management. In L. Palen, M. Büscher, T. Comes, & A. L.

Hughes (Eds.), *12th Proceedings of the International Conference on Information Systems for Crisis Response and Management, Krystiansand, Norway, May 24-27, 2015* (pp. 244–253). ISCRAM Association.

Riechert, T., & Berger, T. (2009). Leveraging semantic data Wikis for distributed requirements elicitation. In *Icse Workshop on Wikis for Software Engineering, 2009: Wikis4se '09 ; 19 - 19 May 2009, Vancouver, Canada ; in conjunction with the 2009 IEEE 31st International Conference on Software Engineering (ICSE 2009), May 16- 24, 2009* (pp. 7–13). IEEE. https://doi.org/10.1109/WIKIS4SE.2009.5069992

Riechert, T., Lauenroth, K., Lehmann, J., & Auer, S. (2007). Towards semantic based requirements engineering. In *Proceedings of the 7th International Conference on Knowledge Management (I-KNOW)* (pp. 144–151).

Robertson, S., & Robertson, J. (2013). *Mastering the requirements process: Getting requirements right* (3rd ed.). Addison-Wesley.

Rolland, C., Achour, C. B., Cauvet, C., Ralyté, J., Sutcliffe, A., Maiden, N., Jarke, M., Haumer, P., Pohl, K., Dubois, E., & others (1998). A proposal for a scenario classification framework. *Requirements Engineering*, *3*(1), 23–47.

Rosson, M. B., & Carroll, J. M [John Millar]. (2008). *Usability engineering: Scenario-based development of human-computer interaction* [Repr., 2. Druck]. *The Morgan Kaufmann series in interactive technologies*. Kaufmann. https://doi.org/10.1016/B978-1-55860-712-5.X5000-0

Rupp, C. (2014). *Requirements-Engineering und -Management.* (6th ed.). Hanser.

Sabatucci, L., Ceccato, M., Marchetto, A., & Susi, A. (2015). Ahab's legs in scenario-based requirements validation: An experiment to study communication mistakes. *Journal of Systems and Software*, *109*, 124–136. https://doi.org/10.1016/j.jss.2015.07.039

Schauwecker, S., Gascón, E., Park, S., Ruiz-Villanueva, V., Schwarb, M., Sempere-Torres, D., Stoffel, M., Vitolo, C., & Rohrer, M. (2019). Anticipating cascading effects of extreme precipitation with pathway schemes - Three case studies from Europe. *Environment International*, *127*, 291–304. https://doi.org/10.1016/j.envint.2019.02.072

Siegemund, K. (2014). *Contributions To Ontology-Driven Requirements Engineering* [Dissertation]. TU Dresden, Dresden.

Sutanto, S. J., van der Weert, M., Wanders, N., Blauhut, V., & van Lanen, H. A. J. (2019). Moving from drought hazard to impact forecasts. *Nature Communications*, *10*(1), 4945. https://doi.org/10.1038/s41467-019-12840-z

Sutcliffe, A. (1998). Scenario-based requirements analysis. *Requirements Engineering*, *3*(1), 48–65. https://doi.org/10.1007/BF02802920

Sutcliffe, A. G., Maiden, N. am, Minocha, S., & Manuel, D. (1998). Supporting scenario-based requirements engineering. *IEEE Transactions on Software Engineering*, *24*(12), 1072–1088.

Terti, G., Ruin, I., Kalas, M., Láng, I., Cangròs i Alonso, A., Sabbatini, T., & Lorini, V. (2019). ANYCaRE: a role-playing game to investigate crisis decision-making and communication challenges in weather-related hazards. *Natural Hazards and Earth System Sciences*, *19*(3), 507–533. https://doi.org/10.5194/nhess-19-507-2019

Touloum, K., Idoughi, D., Seffah, A., & Sabroux, C.-R. (2013). Scenarios + Touchpoints = A Method for Analyzing Crisis Situations and Designing Management and Rescue Services. In T. Comes, F. F. Fiedrich, S. Fortier, J. Geldermann, & T. Müller (Chairs), *ISCRAM*, Baden-Baden.

Tudorache, T. (2006). *Employing Ontologies for an Improved Development Process in Collaborative Engineering* [Dissertation, Technische Universität Berlin, Berlin]. DataCite.

Tuma, K., Calikli, G., & Scandariato, R. (2018). Threat analysis of software systems: A systematic literature review. *Journal of Systems and Software*, *144*, 275–294. https://doi.org/10.1016/j.jss.2018.06.073

Van Lanen, H. A., Vitolo, C., Di Napoli, C., Sutanto, S., D'Andrea, M., Bergman, T., Duo, E., Fernandez Montblanc, T., Gascón, E., Di Giuseppe, F., Láng, I., von Lerber, A., Park, S., Pignonen, F., Roca-Sancho, J., Schauwecker, S., Ciavola, P., Berenguer, M., Koistinen, J., & Llort, X. (2019). A pan-European multi-hazard early warning system: ANYWHERE MH-EWS. *Geophysical Research Abstracts*, *21*.

Weidenhaupt, K., Pohl, K., Jarke, M., & Haumer, P. (1998). Scenarios in system development: current practice. *IEEE Software*, *15*(2), 34–45.

Xi, M., Li, Y., Wei, Y., Wang, N., Yin, Y., Luo, Z., Deng, S., Mao, Y., & Yin, J. (2019). A Scenario-Based Requirement Model for Crossover Healthcare Service. In *2019 IEEE World Congress on Services (SERVICES)* (pp. 252–259). IEEE. https://doi.org/10.1109/SERVICES.2019.00075



ISCRAM 2024
MÜNSTER, GERMANY

Disaster Public Health and Healthcare Informatics

Track Chairs

Reem Abbas, Auckland University of Technology

Zeno Franko, Medical College of Wisconsin

Samanta Varela Castro, Center for Research & Teaching in Economics

Abigael Collier, Medical College of Wisconsin

Julie Dugdale, University Grenoble Alps J

ose J. Gonzalez, University of Agder

Accounting for systemic risk in analysis of crisis scenarios

Jose J. Gonzalez
Department of ICT, Faculty of Engineering and Science, Centre for Integrated Emergency Management (CIEM), University of Agder, Kristiansand, Norway
josejg@uia.no

Reem Abbas
Department of Computer Science and Software Engineering
Auckland University of Technology
Auckland, New Zealand
reem.abbas@aut.ac.nz

Sigurd Paulsen
Dept. of Organization, Crisis Management and Information Security,
Kristiansand municipality, Norway
sigurd.paulsen@kristiansand.kommune.no

ABSTRACT

Systemic risk is ubiquitous in our increasingly globalised world owing to the interconnections within and across sectors. However, a review of National Risk Assessments performed in OECD member countries reveals that systemic risk is rarely mentioned, and it is never accounted for in those assessments.

A powerful method accounts for systemic risk and identifies mitigating strategies using participatory modelling with stakeholders: strategy mapping with *Strategyfinder*. The risk model is represented as a directed graph consisting of risk nodes and directed edges (arrows) expressing causal influences. The method originated in the management of large, complex projects, and it has been applied to systemic risks in critical infrastructures and pandemics.

We demonstrate the *Strategyfinder* method using a scenario describing a cyberattack on electronic communication infrastructure leading to serious consequences for health and lives. The scenario is borrowed from the report Analysis of Risk Scenarios by the Norwegian Directorate for Civil Protection.

Keywords

National Risk Assessments, Risk scenarios, Disaster risk reduction, Vicious cycles, Cascading effects.

INTRODUCTION

A Wider Perspective of Systemic Risk

Systemic risk has been increasingly recognised during the last 2-3 decades as ubiquitous. While the disaster risk reduction community is aware of systemic risk in financial systems (Kim et al., 2022, p2; Schweizer, 2021, p90), the community is less conscious of the role of systemic risk in disruption and delays in complex projects (Gonzalez & Eden, 2023). Not being sufficiently aware of the literature on systemic risk in complex projects (e.g., Ackermann et al. 2007; Ackermann et al. 2011; Williams et al., 1997), the disaster risk reduction community has not yet adopted a validated approach of strategy mapping for disaster risk management using participatory modelling with a team of stakeholders (practitioners, and power-brokers). Using this approach, the team develops risk mitigations strategies deemed effective and practical. Recent evidence of the approach's applicability in disaster risk management has emerged (Abildsnes 2023; Bryson et al., 2023; Eden & Gonzalez, 2023).

CoRe Paper – Disaster Public Health & Healthcare Informatics
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Feedback Loops as Result of Interconnected risks

While there are many definitions of systemic risk, all mention feedback (a.k.a. circular causality) among its characteristics (Centeno et al, 2015; Flood et al., 2022; Gonzalez & Eden, 2023; Kaufmann & Scott, 2003; Kim et al, 2022; Reichstein et al., 2021; Renn et al., 2019; Schwarcz, 2008; Schweizer, 2021; UNDRR & UNU-EHS, 2022).

Systemic risk occurs because of the interconnections between risks. When risks give rise to further risks, as is the case in financial systems, in complex projects, and in many disasters, the consequences become very complicated. In all of them, a risk can cause a plethora of other risks, and significantly that risks reinforce themselves through feedback loops (a.k.a., vicious cycles).

Major disasters, like the COVID pandemic, trigger many vicious cycles across the critical infrastructures of nations (in fact, across the whole world). Cascading effects happen both within sectors and across sectors. The mitigation of systemic risks cannot proceed by mitigating separately single vicious cycles. Rather, the identification of the potency of risks – those risks with the most vicious cycles – is a must, to make those most potent risks the priority for risk mitigation measures (Gonzalez et al., 2021).

UNDRR Report on Systemic Risk Mitigation

Systemic risk has got prominent attention from the United Nations Office for Disaster Risk Reduction (UNDRR) in its two last reports on disaster risk reduction – counting together 730 pages. The Global Assessment Report on Disaster Risk Reduction 2019 (UNDRR, 2019), and the Global Assessment Report on Disaster Risk Reduction 2022 (UNDRR, 2022), are extensive reviews of the emergent science of systemic risk and of the challenges posed by systemic risk. In the following we refer to those reports as GAR2019 and GAR2022 for brevity.

GAR2019 states that systemic risks are triggered by the complexity and the interconnections of our increasingly globalised world, whereby the resulting networks determine the exposure and vulnerability shaping the dynamic interactions among the Sendai Framework, the 2030 Agenda, the Paris Agreement, New Urban Agenda, and the Agenda for Humanity (UNDRR, 2019, p32). Assessment and management methodologies for systemic risks are not yet part of the current operations of the twenty-first century risk management institutions; however, there is a growing sense of urgency for a paradigm shift (ibid, p44). The science of systemic risk and systemic risk management is still in "primordial state" (ibid, p146). GAR2022 asserts that little progress has happened in the three years since the previous global assessment report (UNDRR, 2022, p146).

However, the GAR2022 report fails to capture recent advancements utilising strategy mapping. Several publications (Abildsnes et al., 2023; Gonzalez et al., 2021; Gonzalez & Eden, 2022; Eden & Gonzalez, 2023; Gonzalez & Eden, 2023) and a recent report (Bryson et al., 2023, p24-30) describe the extension and application of the approach which emerged in complex project management to systemic risk in the COVID pandemic.

National Risk Assessments – A Fertile Ground for Systemic Risk Management?

The national civil protection agencies are tasked with maintaining a general overview of risk and vulnerability within society. Systemic risk became a hot topic for disaster risk reduction about 2 -3 decades ago. Thus, it is natural to ask: do the national civil protection agencies account for systemic risk in their national risk assessments? If the answer is 'not yet', is there a fertile ground in future national risk assessments for using the recent advancements in strategy mapping (ibid)?

Our method to answer the question consists of two steps:

1. Check if a representative report on national risks assessments proves that systemic risk is being accounted for.
2. If the answer to #1 is 'no', apply strategy mapping to an existing national risk assessment to account for systemic risk.

We proceed to these steps in the following two sections.

IS SYSTEMIC RISK ACCOUNTED FOR IN NATIONAL RISK ASSESSMENTS?

For the first step, we examine the report on disaster risk analysis procedures by the Organisation for Economic Co-operation and Development (OECD). The OECD report presents a cross country perspective of national risk assessments including governance practices for National Risk Assessments and the assessment results in the 20 OECD countries including Australia, Austria, Canada, Denmark, Estonia, Finland, Germa ny, Hungary, Korea, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland,

United Kingdom, and the United States (OECD, 2018).

The OECD report states: "National Risk Assessments are used to support risk management decisions in a rapidly changing global risk landscape characterized by increasingly complex, interconnected societies and highly mobile people, information and goods." In increasingly complex and interconnected societies the risks become increasingly systemic.

Among the countries reviewed in the OECD report, New Zealand is the only country that included systemic risk in its key definitions, stating that systemic risks are those that "arise from the interactions between interdependent elements of the economy and society". One of the criteria used in New Zealand to characterise risks requiring a national response relates to have "multiple or inter-related problems which when taken together constitute a national or systemic risk".

However, none of the national risk assessments reviewed in (OECD, 2018) report systemic risks.

APPLYING STRATEGY MAPPING TO AN EXISTING NATIONAL RISK ASSESSMENT

For the second step, we examine if a 'promising' risk scenario implicitly accounts for systemic risk. By a 'promising' risk scenario we mean a very detailed scenario where risks cause cascading effects across several critical infrastructures.

The Norwegian Directorate for Civil Protection has been conducting national analyses of risk scenarios since 2011. The risk analyses take a social science approach and are based on qualitative data, expert assessments, and broad participation in the analysis processes. The analyses, which are updated regularly, are very detailed.

Among the 25 risk scenarios described by the Norwegian Directorate for Civil Protection in their last issue of Analysis of Crisis Scenarios (Norwegian Directorate for Civil Protection, 2020), we select a scenario which *a priori* could be expected to contain many interconnected risks, viz. scenario 18.2 "Cyber attack on electronic communications infrastructure" (Norwegian Directorate for Civil Protection, 2020, p204 -207). The scenario describes a cyber attack against the central nodes in the transport network of the Norwegian telecommunications company Telenor. According to the national risk analysis, the attack would cause all electronic communication devices in Norway to fail completely for five days, with massive consequences on critical infrastructures, particularly on life and health.

In the scenario, Telenor's transport network for electronic communications is described as the only nationwide in Norway. The situation has changed since. But although two additional networks, Broadnet and Telia, increased its coverage since the report was produced, Telenor is still the largest transport network in Norway, with more than 200,000 kilometres of fibre and copper cables, and more than 10,000 base stations.

Therefore, the scenario 18.2 "Cyber attack on electronic communications infrastructure" described in the report by the Norwegian Directorate for Civil Protection (ibid, p204-207) is still suitable for answering the question whether the social science-based scenario analysis employed by the Norwegian Directorate for Civil Protection considers systemic risk.

Method

Reinforcing feedback loops of risks (a.k.a. vicious cycles) are the cause of escalating cascading effects. They must be expected in a massive cyber attack to the electronic communication infrastructure. Accordingly, we check whether the risk analysis of scenario 18.2 in (Norwegian Directorate for Civil Protection, 2020, p204 -207) contains feedback loops. To be able to check whether the risk analysis of scenario 18.2 contains feedback loops we must first express the scenario as a formal model capable of being analysed for the existence of feedback loops.

Such formal model can be created with *Strategyfinder*™. *Strategyfinder* is a browser-based professional version of methods and tools developed at the University of Strathclyde since the 1990'ies. With *Strategyfinder* teams collaboratively work in person or virtually on messy problems, to develop strategies, and manage risks. However, in this paper we use *Strategyfinder* to create a formal risk model by reverse engineering scenario 18.2.

A *Strategyfinder* model of risks is a mathematical object known as a directed graph. In a directed graph there are nodes and directed edges ('arrows') expressing causality (i.e., that a node impacts causally on another node).

The nodes in a *Strategyfinder* risk model are risk statements: sentences of typically around 6-12 words describing risks. We adopt Lupton's (2013) definition of a risk: "a phenomenon that has the potential to deliver substantial harm, whether or not the probability of this harm eventuating is estimable".

The first step to create a *Strategyfinder* version of the scenario is parsingthe description of scenario 18.2 to identify

the risks. Since *Strategyfinder* assigns a reference number to each statement, we number the risks in the order in which they are parsed in the subsections below.

The risks are listed in the subsections of the scenario 18.2 "Assessment of vulnerability", "Assessment of likelihood", "Assessment of consequences – Life and health", "Assessment of consequences – Economy", "Assessment of consequences – Social stability", and "Democratic values and capacity to govern". There is some redundancy in the scenario description: some risks are mentioned several times in different subsections. Hence, some statements which are obvious risks appear below without reference number (meaning that they are present in the *Strategyfinder* version using a slightly different wording).

In the following subsections we parse the text from scenario 18.2 (ibid, 204-207), labelling the risks with numbers in the same order as they appear in the scenario. In other words, we proceed by copy/pasting from scenario 18.2 and adding sequent numbers to each risk.

Parsed Risks from Scenario 18.2, Subsection "Assessment of Vulnerability"

[1] Telenor's transport network for electronic communications is the only nationwide one in Norway. [2] Broadnet has a transport network that covers 90 Norwegian towns and cities. However, [3] some of the infrastructure is shared with Telenor, and a [4] failure in Telenor's network would therefore also [5] knock out Broadnet's network. [6] Other electronic communications operators such as Telia, Ice, the emergency services' network (Nødnett) and others, mainly use Telenor's transmission infrastructure, and to some extent Broadnet's. [7] National radio and television companies rely on Telenor's transport network to get signals to transmitters. [8] Telenor's transport network is robust and well protected. However, in the event of [9] any network outage there are no alternatives. [10] Satellite telephony and radio communications do not have nearly enough capacity to meet the need for communication.

The most serious consequences are: [11] Management of the crisis at a political and administrative level becomes difficult, with [12] reduced opportunities for communication and coordination. The [13] disappearance of radio, television and the Internet means that [14] important information channels for the public would be unavailable. [15] Rescue efforts become more difficult in the [16] absence of telephony and when the [17] functionality of emergency numbers and the emergency services' network (Nødnett) is severely hampered. [18] Railway and air traffic would stop. [19] Road and maritime traffic would also experience problems. [20] Failures in payment services would cause [21] major challenges for the commercial sector and the public.

Parsed Risks from Scenario 18.2, Subsection "Assessment of Likelihood"

In this subsection of scenario 18.2 some of the statements are risk attributes, rather than risks themselves. We opt for considering them in the model, since the attributes have causal implications on the risks.

[22] Conducting a successful cyber attack as outlined in this scenario [23] requires a very high level of expertise and capacity – also with regard to intelligence. It is believed that only a [24] few actors have such capabilities. The likelihood would depend on the [25] international threat picture. The Norwegian Police Security Service's open threat assessment for 2018 states that "[26] enterprises within the Norwegian defence and public security sector, public administration, research and development and critical infrastructure are assessed to be particularly at risk of becoming intelligence targets." This applies with respect to [27] both network operations and more traditional intelligence gathering. [28] Several countries' intelligence services have interests within these areas.

Parsed Risks from Scenario 18.2, Subsection "Assessment of Consequences – Life and Health"

[29] Reduced opportunity to notify the emergency services in the event of acute events, [30] no possibility to call an ambulance in the normal manner, [31] inadequate communication and coordination between the emergency services, as well as [32] reduced efficiency and delayed treatment of patients in the health and care sector will have [33] consequences for life and health.

The analysis is based on an assumption that around 5% of acutely sick or injured people (who would otherwise have survived) would die. This means that, overall, the scenario would result in around 10 more deaths per day, or approximately 50 deaths in a five-day period, which represents [34] an increase of around 10% in relation to the normal daily mortality rate. [35] A number of planned treatments would be cancelled due to reduced efficiency. [36] During a five-day period, we estimate that 200–300 people would become significantly more ill due to reductions in the provision of treatment.

Parsed Risks from Scenario 18.2, Subsection "Assessment of Consequences – Economy"

It is assumed that the [37] direct economic losses will be NOK 2 to 10 billion and that they will mainly be associated with [38] the necessary repair and replacement of physical components and infrastructure.

The indirect economic losses will be associated with [39] loss of income, production losses and a decline in consumption, orders, and deliveries. A functioning payment system is a prerequisite for being able to pay for the delivery of goods and services, as well as the trading of financial instruments. Around a third of the normal production, or approximately NOK 13 billion, will be lost as a result of the loss of electronic communications. Even if some of the lost revenue can be recouped, it is assumed that [40] the net loss will exceed NOK 10 billion.

Parsed Risks from Scenario 18.2, Subsection "Assessment of Consequences – Social Stability"

The scenario would result in [41] significant reactions among the public in the form of anxiety, insecurity, fear and a feeling of powerlessness. The incident would be experienced as sudden and unfamiliar. A lack of information would significantly contribute to the level of anxiety. The reactions would be reinforced by [42] people involved in emergency situations being unable to get through on emergency numbers.

Parsed Risks from Scenario 18.2, Subsection "Democratic Values and Capacity to Govern"

The outage of electronic communications would cause [43] major challenges for governance and crisis management. The event would be perceived as a significant violation of shared cultural and democratic values, as well as basic individual rights and personal safety.

STRATEGYFINDER MODEL OF SCENARIO 18.2

The *Strategyfinder* modelin this section sticks to the description of the risk scenario 18.2 provided in (Norwegian Directorate for Civil Protection, 2020, p204 -207). Accordingly, from now on we will refer to the corresponding *Strategyfinder* model as the 'basic risk model'.

Fig. 2 on p7 displays the *Strategyfinder* basic risk model of the scenario 18.2. Explaining all the risk nodes and causal relations (represented by directed links, i.e., arrows) would be extensive. Hence, we use *pars pro toto* a restricted view produced with the *Strategyfinder* analysis tool 'dependency range' applied to #11 'management of the crisis at a political and administrative level becomes difficult' (Fig. 1 on p6). Fig. 1 shows only the risks and issues that directly or indirectly impact on #11.

First, notice that all the risks and issues shown in Fig. 1 are taken from the list of risks and issues parsed from scenario 18.2, cf. p4-5. Then, notice that some items, e.g., #9, #11, and #22, are formatted differently from the others. The reason for the different formatting will be explained below.

The arrow #12→#11 expresses that reduced opportunities for communication and coordination causes that the management of the crisis at a political and administrative level becomes difficult. The arrow #15 →#12 expresses that absence of telephony causes reduced opportunities of communication and coordination. Hence, #15 impacts indirectly on #11. In addition to absence of telephony, the fact that in the event of an outage to the transport network there would be no alternatives means that #9 impacts directly causally on #11 in addition to doing it indirectly via #15. Thus, one has a direct causal impact on #11 expressed by the arrow #9 →#11, and an indirect causal impact on #11 expressed by the series of arrows #9 →#15→12→#11.

If there is an outage, risk #9 (no alternatives for electronic transport in case of an outage) causes #13 , disappearance of radio, television, and the internet, hence the arrow #9 →#13; crucial for that is the impact #7→#13, viz., that radio and television companies rely on Telenor's transport network to get signals to transmitters.

Multiple causes impact risk #9 ('no alternatives for electronic transport in case of an outage'):

- #4→#9, since failure in the Telenor's transport network directly impacts on #9, and because of #1→#9, that Telenor's transport network is the only nationwide one in Norway;
- #5→#9, since Broadnet, who could provide part of the electronic transport, is impacted #4→#5 from the failure in Telenor's transport network;
- #6→#9 expresses that other alternatives depend on Telenor or on Broadnet. Thus, no alternative operates in the case of the outage that knocks out Telenor and Broadnet.

The arrow #8→#4 is labelled with a red minus sign. Arrows with a red minus sign express *negative polarity*, viz., that *increasing* the cause *reduces* the causal impact. Here it means that the high robustness and degree of protection

of Telenor' network decreases the probability that a network failure happens. Arrows with positive polarity are in *Strategyfinder* not labelled with a + sign. I.e., no polarity sign means positive polarity by default.

Consider now #22 'conducting a successful cyber attack'. Risk #22 represents the desired outcome of the malicious agent. Hence, #22 has been formatted using the predefined *Strategyfinder* category 'STRATEGY'.

Conducting a successful cyber attack leads to failure of Telenor's network, hence the causal link #22 →#4.

Conducting a successful attack requires network operations and intelligence gathering by the malicious agent, expressed by #27→#22.

We skip explaining more arrows, since at this stage they should be self-explanatory.

The outcome #11 (the management of the crisis becomes difficult) is a goal for the agents behind the cyber attack. Hence, it has been formatted using the predefined *Strategyfinder* category 'GOAL'. Goals are outcomes desired on their own right. Goals have mostly only incoming arrows. Nevertheless, one can have goals (desired outcomes on their own right) one or two levels below a final goal. E.g., #12 ('reduced opportunities for communication and coordination in Norway') is arguably also a goal for the malicious agent. However, it is also a very busy node, with many ingoing and outgoing arrows, in the risk model. Hence #12 was rather formatted with the FOCUS category (red font), as were #9 and #32 for the same token (very busy nodes).

Strategyfinder has analysis tools to detect the busiest elements in the risk model (e.g., using the analysis submenu 'counting ingoing and outgoing links') and to detect 'heads' (elements with only ingoing arrows). The nodes #11, #34, #37, #40 and #41 are all 'heads' and have been formatted with the *Strategyfinder* category GOAL. Note that they are **negative** goals, i.e., intended outcomes of the cyber attack, which need mitigation strategies and actions on the side of the defenders (Eden, 2004; Eden & Gonzalez, 2023).

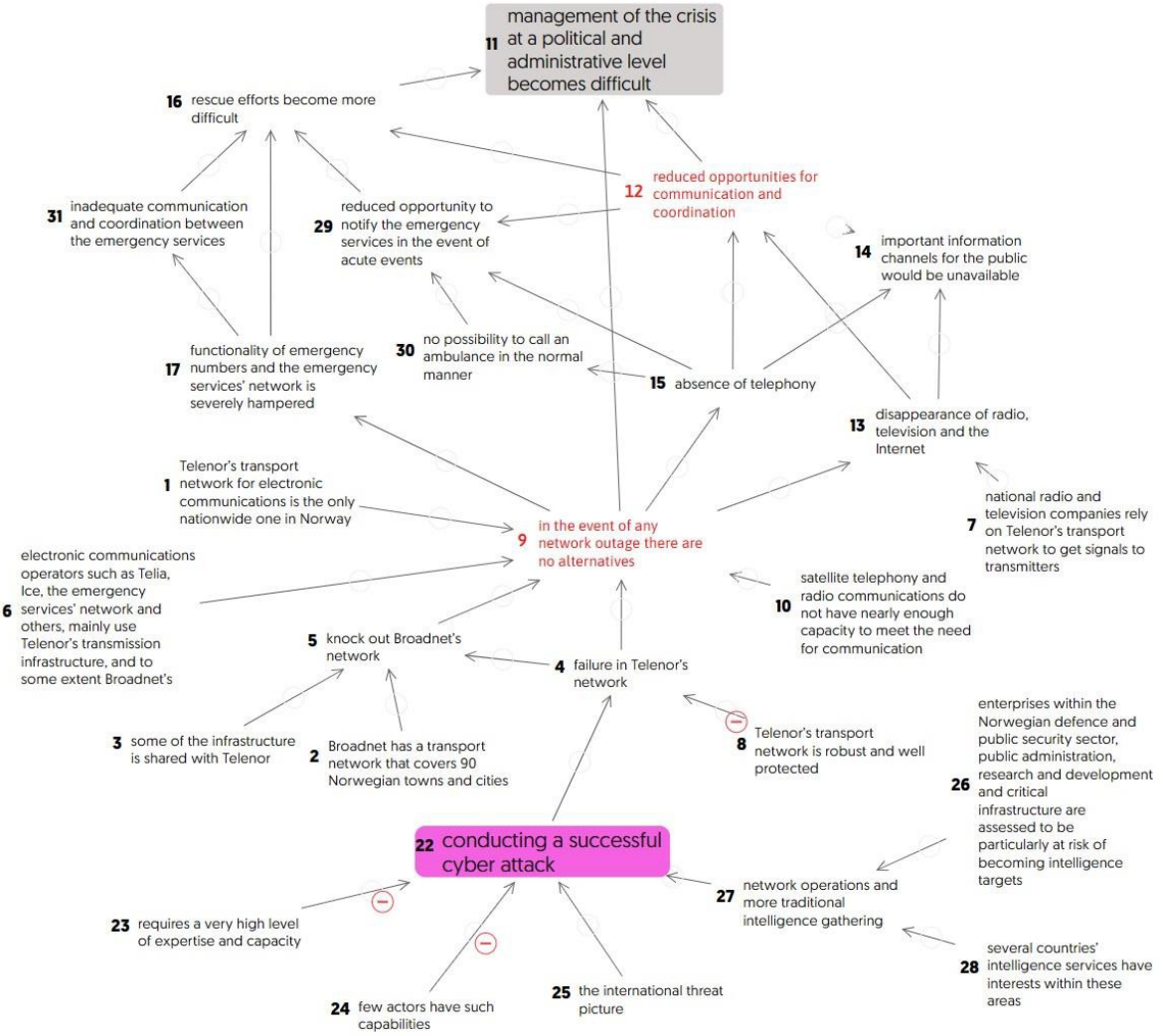


Figure 1 A restricted view of the basic risk model (shown in full on Figure 2) with the risks and issues that directly or indirectly impact on #11 'management of the crisis at a political and administrative level becomes difficult'.

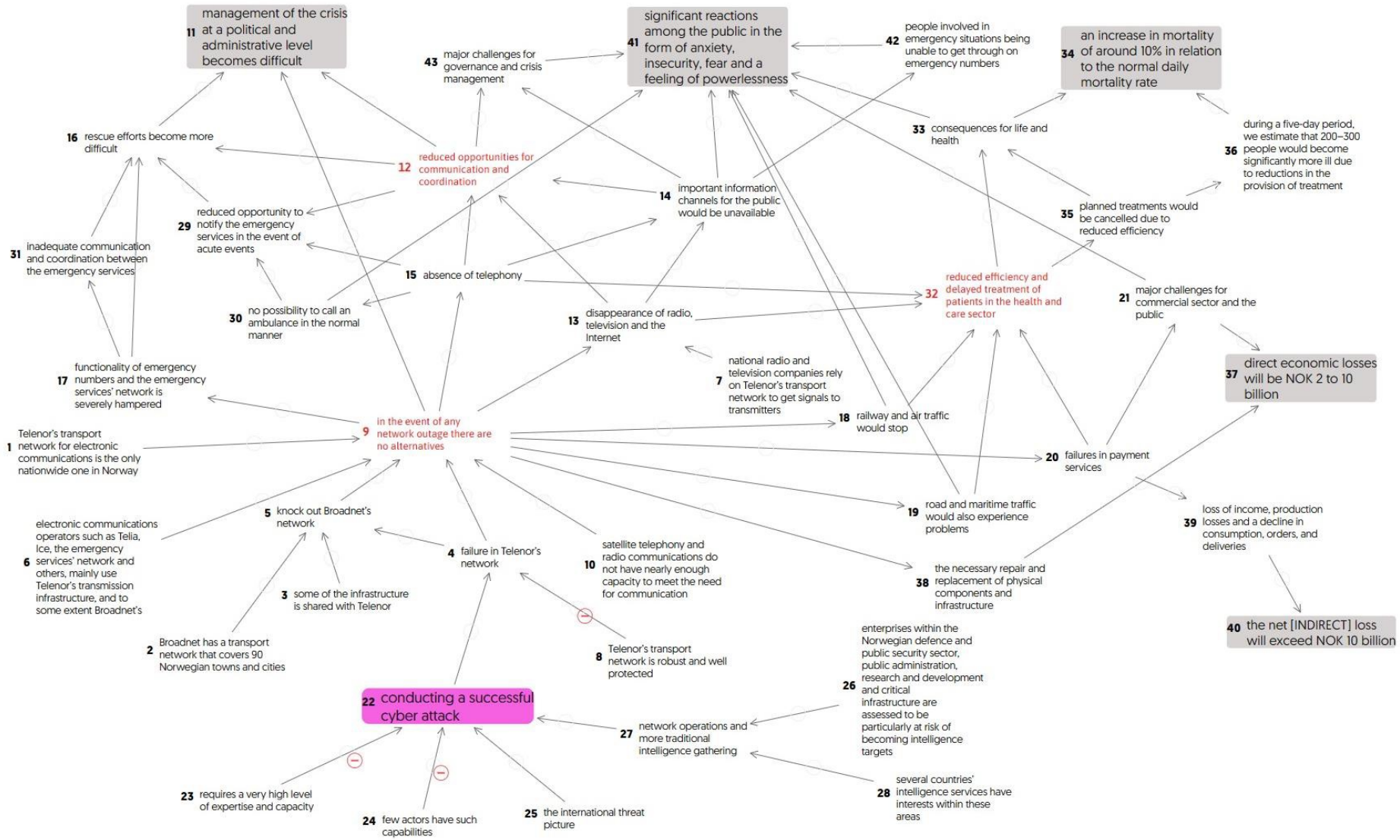


Figure 2 The complete risk model corresponding to Scenario 18.2.

Does the Basic Risk Model Exhibit Systemic Risk?

Systemic risk has a particular kind of complexity, viz., *dynamic complexity*. Dynamic complexity arises if there is causal feedback in a system (Sterman, 2000 , p12-13, 138-141). In the case of disasters, one has reinforcing feedback of risks (vicious cycles of risks). Systems with vicious cycles have dynamic complexity, which makes them very difficult to manage (Eden & Gonzalez, 2003; Gonzalez & Eden, 2023).

A powerful feature of *Strategyfinder* is its ability to detect and identify feedback loops using analysis algorithms applied to the directed graph representing the risk model. Using this feature on the basic risk scenario model, *Strategyfinder* reported 'no loops detected'. Accordingly, the model of risks and their causal interconnections matching the description provided in (Norwegian Directorate for Civil Protection, 2020, p204 -207) must be incomplete, since it does not have feedback loops and, hence, lacks major cascading effects. A massive cyber attack leading to a full failure of the electronic communication devices for five days in Norway must be expected to escalate the risks over time, i.e., to generate vicious cycles of risks (i.e., reinforcing feedback loops of risks and risk outcomes).

ACCOUNTING FOR SYSTEMIC RISK USING THE *STRATEGYFINDER* METHOD

We recall from the previous section, that scenario 18.2 described in (Norwegian Directorate for Civil Protection, 2020, p204-207) does not account for all the causal links among the listed risks, since the basic risk model does not have feedback loops. Without feedback loops the dynamic behaviour of the basic risk model appears to be far too simple. There is no escalation of risks over time. Such escalation must be expected as consequence of a massive cyber attack with cascading effects on several critical infrastructures.

In addition to the missing causal links, which must be expected within the critical infrastructures mentioned in (ibid, p204-207), viz. electronic communication, life and health, transport and finances, the description also seems to be incomplete regarding risks content. For instance, a massive cyber attack disrupting all electronic communication networks for five days would also knock out energy production, in particular hydroelectric energy production (Norway's main source of energy for the public, for much of its industry and for hospitals). Among the consequences of energy disruption, a major one would be the impact on livestock (cattle, pigs, etc), and fish farming (one of the main sources of Norway's prosperity), as well as the impact on refrigerators (household, commerce, and industry).

Extending the risk model to consider all aspects of a massive cyber attack is beyond the scope of this paper. (We mention in passing that such a full risk model would be needed in order to develop adequate mitigating strategies of the risks.) For the purpose of this paper, it is sufficient to show that a modest enhancement of the basic risk model generates vicious cycles of risks, with a resulting dynamic behaviour of risks escalation. In fact, we show that the escalation causes both inner-sector and inter-sector cascading effects.

Identification of Missing Causal Links and Their Consequences

Enhancing the basic risk model by adding a few more obvious causal links among the risks described in Scenario 18.2 'Cyber attack on electronic communications infrastructure' (Norwegian Directorate for Civil Protection, 2020, p204-207) generates several vicious cycles of risks.

To enhance the basic risk model, the third author of this paper, the crisis manager of the Kristiansand municipality, Norway, added causal links to the risk model. He has ample experience with *Strategyfinder* as he was lead participant for the Kristiansand municipality in the Systemic Pandemic Risk Management project 2020-2023 (Bryson et al., p24-30). The project used *Strategyfinder* and promoted a significant extension of the capabilities of the software. As champion of the approach, after the completion of the Systemic Pandemic Risk Management project, the crisis manager of the Kristiansand convinced the municipality to adopt *Strategyfinder* for systemic risk assessment and management.

On Fig. 3 the added causal links are shown as red arrows. The rational for the added causal links follows on p10:

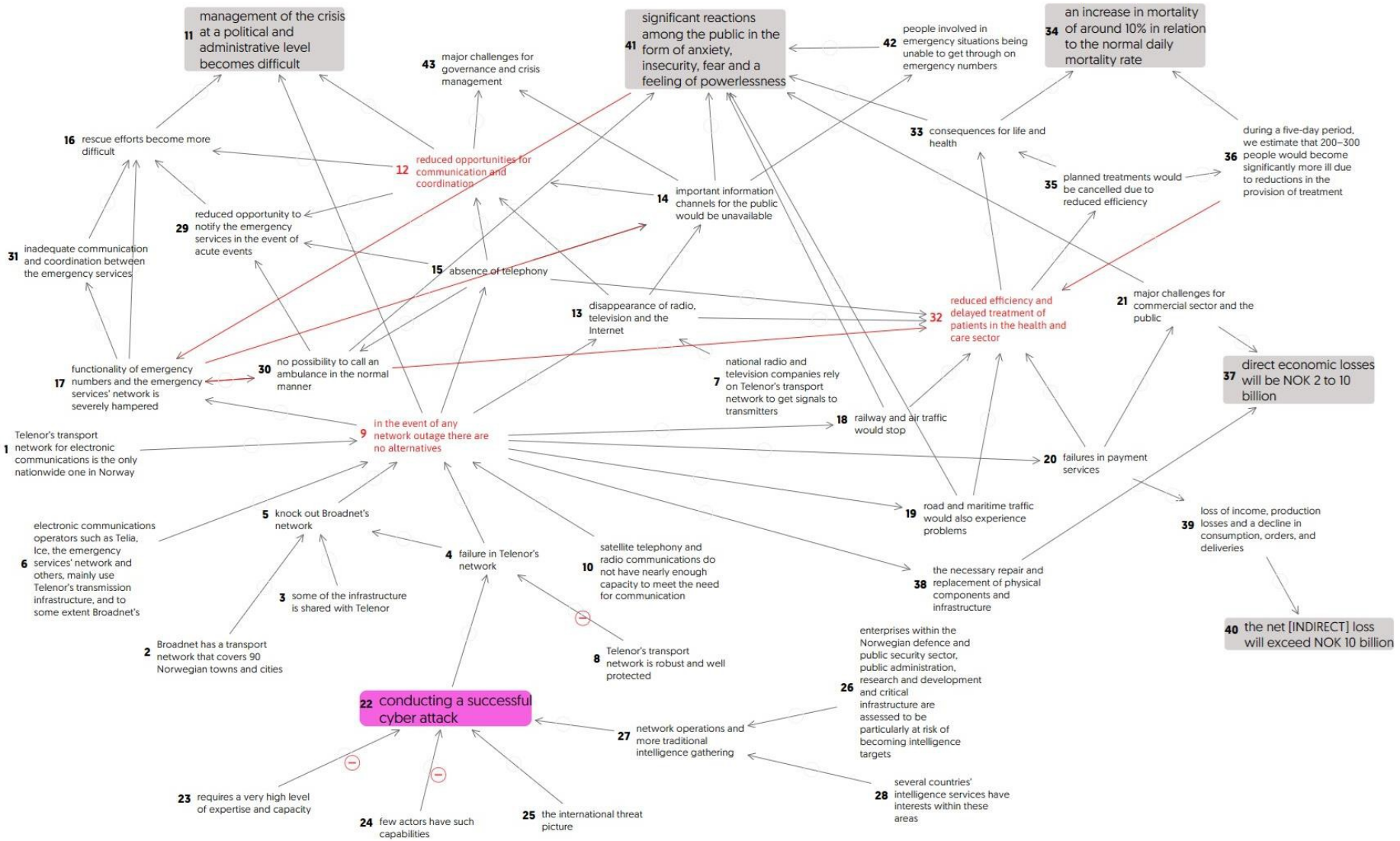


Figure 3 Risk model with added causal links (shown as read arrows). See the main text for the description of the added causal links.

ADDED CAUSAL LINKS AND THEIR EXPLANATION

- #17→30: The functionality of the rescue networks being severely hampered impacts on the ability to call an ambulance in the normal manner.
- #30→17: Not being able to call ambulances in the normal manner adds to the inconveniences hampering the functionality of rescue networks.
- #41→17: Significant reactions among the public (anxiety, fear, hopelessness) is likely to exacerbate conditions that may need request of rescue (which, lacking normal ways of communication, would often be delivered via messengers demanding attention from pressed rescuers).
- #30→32: No possibility to call ambulances impacts on the efficiency of treatment of patients and adds delays.
- #17→14: Since rescue services may detect conditions that demand alerting the public.
- #36→32: A significant increase in serious illnesses will reduce the efficiency treatments in hospitals and cause time delays

In the resulting risk model (Fig. 3), *Strategyfinder* detects 7 feedback loops, all acting as vicious cycles. Fig. 4 shows the result of the analysis.

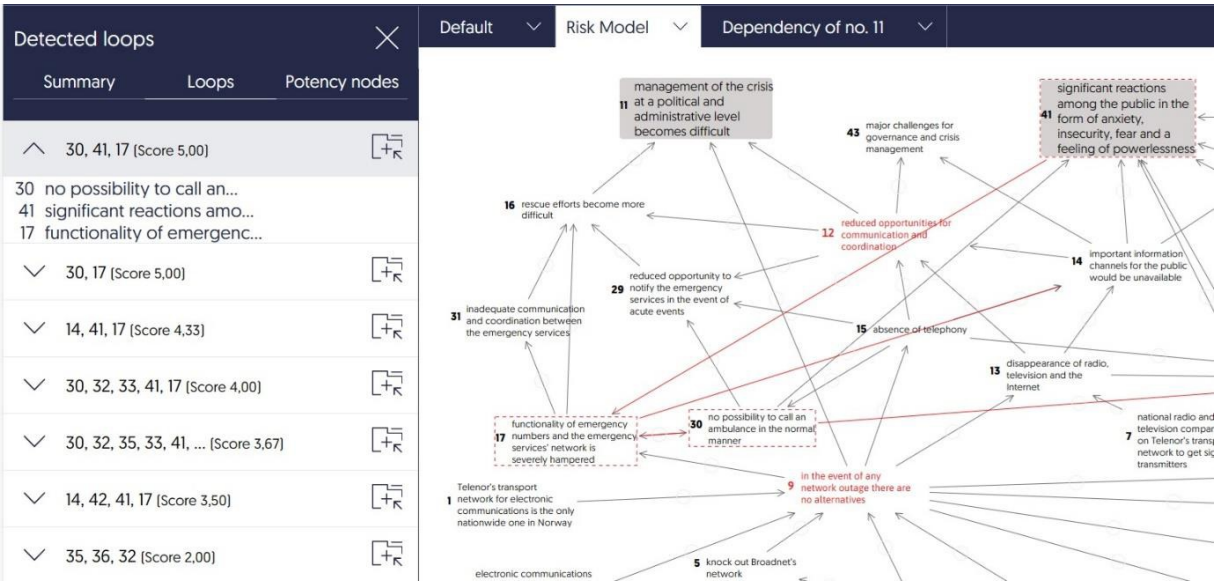


Figure 4 Detected feedback loops. Clicking on a detected loop (here, 17→30→41→17), selects the corresponding risk nodes on the risk model (shown in red colour framed with dashed lines).

We skip the explanation of the scores, since they are not relevant for the purpose of this paper (other than to comment that higher scores mean more potency of the risks).

The symbol to the right of the score indicates that clicking on it creates a new view and the view displays the corresponding loop.

Figure 5 shows the new view (labelled 'Loop 30,17') and the loop (vicious cycle) #17↔#30; note that #17 impacts on #30, and #30 impacts back on #17.

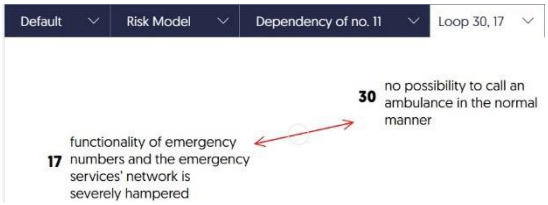


Figure 5 The loop #17↔#30. See main text.

Risk #17 belongs to the electronic communication infrastructure, whereas risk #30 belongs to the public health infrastructure. In other words, the vicious cycle #17↔#30 causes inter-sector cascading effects.

This loop #17↔#30 is part of 4 vicious cycles and the risk #17 participates as node in 6 vicious cycles.

Figure 6 is a composite figure where the 7 feedback loops detected by *Strategyfinder* in the enhanced risk model are shown.

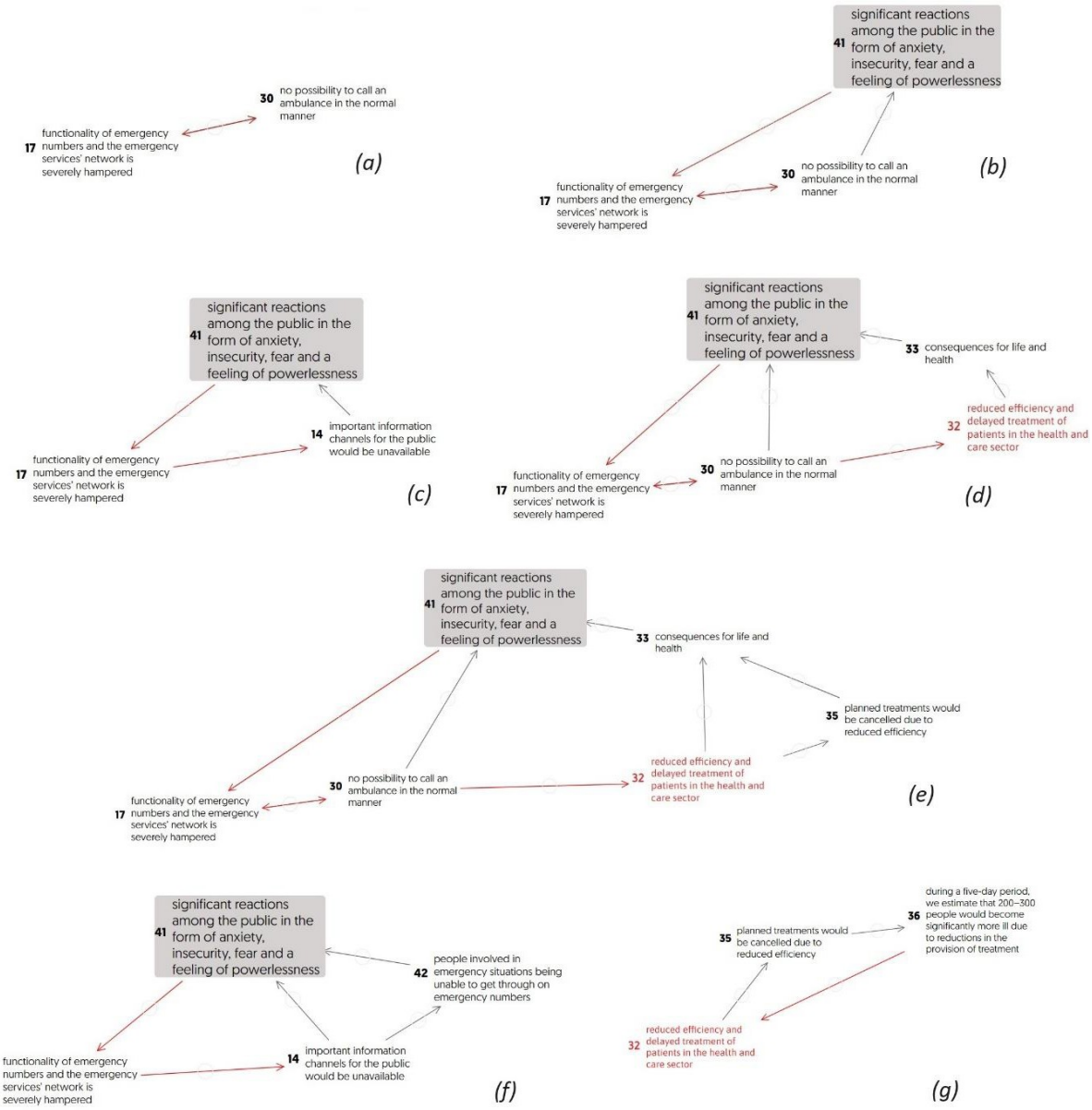


Figure 6 The seven feedback loops in the enhanced risk model, each loop acting as a vicious cycle.

The loop labelled (a) on Figure 6 is identical to the loop #17↔#30. It was discussed in the previous page.

The loop labelled (b) is a nested loop, consisting of two single loops, viz. #30 ↔#17 and #30→#41→#17→#30, which share the node #17. In a nested loop, each of its components act as a single vicious cycle and the interaction between the single loops potentiates their impact. As in the case of the vicious cycle labelled (a), the vicious cycle labelled (b) causes inter-sector cascading effects. As mentioned earlier, risk #41 is likely to generate riots, affecting yet another critical infrastructure (security services). However, the scenario 18.2 does not mention this likely consequence.

The loop labelled (c) is the single loop #14→#41→#17→#14. This loop causes inter-sector cascading effects.

The loop labelled (d) is a nested loop, consisting of three single loops, viz. #30↔#17, #30→#41→#17→#30, and #32→#33→#41→#17→#30→#32. Alternatively, it could be described as consisting of the nested loop labelled (b) on Figure 6, and the single loop #32→#33→#41→#17→#30→#32. This loop causes inner-sector and inter-sector cascading effects.

The loop labelled (e) is also a nested loop. For simplicity, it can be described as consisting of the nested loop labelled (d) – which is itself nested – and the single loop #35→#33→#41→#17→#30→#32→#35. This loop causes inner-sector and inter-sector cascading effects.

The loop labelled (f) is also a nested loop. It consists of two single loops, viz., #17 →#14→#41→#17 and #17→#14→#42→#41→#17. This loop causes inner-sector and inter-sector cascading effects.

Finally, the loop labelled (g) is a single loop: #32 →#35→#36→#32. This loop causes inner-sector cascading effects.

Consequences for Mitigation Strategies

Unmitigated vicious cycles escalate the risks over time and increase the impact of cascading effects. Adding only five more obvious causal links to the basic risk model already generated 7 vicious cycles. If the basic model had been extended by adding more risks (e.g., risks belonging to the energy, the public security infrastructures and food) the number of vicious cycles would have risen significantly. A full extension of the basic model is out of scope. Notwithstanding, using the example below we can explain the principle for how to prioritise targets for mitigating strategies.

The straightforward answer is to target those risks which are in the most vicious cycles, the most potent risks. Figuratively speaking, killing as many birds as possible with one stone (Gonzalez et al., 2021).

Figure 7 displays the outcome of the *Strategyfinder* loop analysis.

The 'nodes' (the risks) in the model are evaluated regarding how many loops (here, vicious cycles) that would be killed by a successful mitigation strategy targeting the node (the 'potency' of the node).

The risk #17 has the highest potency, viz., 6.

Indeed, a glance at Figure 6 shows that #17 is in 6 of the 7 loops. Thus, a strategy mitigating that the cyber attack hampers the functionality of the emergency numbers and the emergency services will counteract the escalation driven by 6 vicious cycles.

Next, the risk #41 has potency 5.

Indeed, a glance at Figure 6 shows that #41 is in 5 of the 7 loops. Thus, a strategy mitigating that the cyber attack causes significant reactions among the public will counteract the escalation driven by 5 vicious cycles.

The reader can compare Figure 7 with Figure 6 and check that the potency of each risk appearing on Figure 7 always corresponds to the number of times the risk is part of a vicious cycle.

The *Strategyfinder* analysis tool reveals much more useful information for strategy development, such as dependencies of a node on other nodes (Fig. 1 is such an outcome), busy nodes (such as #9, #12 and #32) or negative goals (viz., #11, #34, #37, #40 and #41). As to negative goals, cf. p6. The interested reader finds the freely available manual 'Analysis Tools' on https://www.strategyfinder.com/.



Potency	#	Statement
6	17	functionality of emergency numbers and the emergency services' network is severely hampered
5	41	significant reactions among the public in the form of anxiety, insecurity, fear and a feeling of powerlessness
4	30	no possibility to call an ambulance in the normal manner
3	32	reduced efficiency and delayed treatment of patients in the health and care sector
2	33	consequences for life and health
2	35	planned treatments would be cancelled due to reduced efficiency
2	14	important information channels for the public would be unavailable
1	42	people involved in emergency situations being unable to get through on emergency numbers
1	36	during a five-day period, we estimate that 200–300 people would become significantly more ill due to reductions in the provision of treatment

Figure 7 The Strategfinder feedback loop analysis reports the potency of the risks (see main text).

DISCUSSION AND CONCLUSION

The 2019 Global Assessment Report on Disaster Risk Reduction strongly emphasizes the pressing need for a fundamental shift in approaches to managing systemic risk (UNDRR, 2019, p.44). A significant ongoing challenge remains in ensuring that research is not only conducted but also effectively utilized and applied (Boaz & Hayden, 2002).

A recent review by Oliver et al. (2022) focuses on identifying existing activities that engage with research and policy, as well as assessing the impact of these activities on both research outcomes and decision -making processes. The review concludes that many researchers have yet to successfully bridge the gap between scientific knowledge and practical application, resulting in what is described as a substantial and growing volume of unfocused activity, lacking in effectiveness (Oliver et al., 2022, p.704).

Similarly, a recent article by Reichstein et al. (2021) in the prestigious journal Nature emphasizes the importance of creating models that are not only comprehensible but also directly applicable to policymaking. Advocacy for research utilization in practical settings is deemed essential (Boaz & Hayden, 2002, p.440).

The research presented in this paper set out to investigate whether national civil protection agencies account for systemic risk in their national risk assessments. To the best of our knowledge, the approach used in our paper, strategy mapping, is the only one that demonstrably has achieved useable and used mitigation strategies of systemic risk (Gonzalez et al., 2021; Gonzalez & Eden, 2022; Abildsnes et al., 2023; Gonzalez & Eden, 2023; Eden & Gonzalez, 2023; Bryson et al., 2023). The approach uses participatory modelling involving stakeholders, mostly practitioners, experts from different relevant disciplines, and – very important for the implementation of the mitigation strategies – power brokers.

Surprisingly, two decades into the realisation of the systemic risk notion by the disaster risk reduction community, no evidence was found that systemic risk is accounted for in the national risk analyses presented in the OECD countries report (OECD, 2018). From a national security perspective, this might potentially represent a deliberate exclusion by national governments if they deem it sensitive to disclose a thorough analysis of national risks .

Nevertheless, the aim of conducting risk assessments by national civil protection agencies is arguably to 1) provide overviews to politicians and business leaders; 2) communicate risks to a wide audience; and 3) as input for exercises and emergency planning at a regional level. We believe that the importance of systemic risk is too significant to justify its omission in national risks assessments. Emergency planning at the municipal and regional level, where the first line of disaster response will be effectuated, must be detailed and specific. Considering that regional risk assessments must involve sectors with critical infrastructure according to their presence and role, omitting systemic risk assessment at the highest level (national) would hardly encoura ge consideration of systemic risk at the regional level.

The OECD report was published in 2018 and there is a possibility that systemic risk since has become an important part of national risks assessments in some of the countries covered by the OECD report. However, we doubt that systemic risk is sufficiently accounted for in national risk assessments. Our doubt is based on two facts.

First, the most recent report from the United Nations Office for Disaster Risk Reduction (GAR, 2022) acknowledges (ibid, p146) that little progress has been done in this regard over the three years period following the previous global assessment report (GAR, 2019). Arguably, if national risk assessments had employed a novel method to account for systemic risks, the GAR2022 authors would have noticed and reported it.

Second, Norway has not yet conducted *national* risk assessments accounting for systemic risk despite acknowledging its importance and the availability of relevant methods. The Corona Commission established by the Norwegian government in response to the COVID pandemic released a report (The Norwegian Corona Commission, 2020) stating that “In its emergency preparedness efforts, the Government has paid little attention to how risk in one sector is affected by risks in other sectors. A crisis preparedness system in which each sector evaluates its own risks and vulnerabilities, will fail if no one takes responsibility for evaluating the sum of the consequences for society at large. There is a need for a cross-sectoral system that can accommodate the interaction of risks across all sectors. This is a lesson applicable to preparedness in general”.

Systemic risk assessment *has* been developed and applied in Norway recently, but then at the *regional level* in the Systemic Pandemic Risk Management project 2020 -2023 using strategy mapping with *Strategyfinder*. The Systemic Pandemic Risk Management project included partners from the municipality of Kristiansand and the regional hospital (Bryson et al., p24-30), as well as international partners. Carefully selected stakeholders (both experts and decision-makers) participated in online workshops to create, analyse, and validate the risk model, to finally propose effective and practical mitigation strategies and identify roles to monitor and implem ent the strategies (Abildsnes et al., 2023, Bryson et al., 2023). Nonetheless, the Norwegian national authorities have not

yet adopted the method developed and applied at the regional level (in the Systemic Pandemic Risk Management project) for national risk assessments and mitigations strategies.

We mentioned several times that the purpose of strategy mapping is to develop strategies. But answering our research question did not require developing mitigation strategies. Our research question, whether national civil protection agencies account for systemic risk in their national risk assessments, only required examining 1) whether systemic risk was explicitly accounted for in the national reviews described in (OECD, 2018); 2) if there was evidence to that effect after 2018; 3) if there was evidence of feedback loops in a highly promising risk scenario 18.2 provided by the Norwegian Directorate for Civil Protection (2020).

The development of risk mitigation strategies for scenario 18.2 would have demanded a considerable extension of the risk model, i.e., it was out of scope.

On p2 we posed the question whether there is a fertile ground in future national risk assessments for using the recent advancements in strategy mapping. We hope to have created awareness that the method of strategy mapping for assessing and managing systemic risk originating in the field of complex project management deserve being adopted by the disaster risk reduction community.

Hence, we believe that the national civil protection agencies could significantly improve their risk analyses using strategy mapping to account for systemic risk. The report from the IBM Center for the Business of Government reviews strategy mapping (Bryson et al., 2023, p8-15, p41 and p43-46). The report also contains a section on the application of the Systemic Pandemic Risk Management project on mitigation strategies of systemic pandemic risk (ibid, p24-30). Finally, the report has an extensive comparison of different methods and tools for strategy mapping (ibid, p49-53).

CLARIFICATION

None of the authors has ownership in the *Strategyfinder* software and none of the authors is employed by the company owning the software.

REFERENCES

Abildsnes, E., Paulsen, S., & Gonzalez, J. J. (2023). Improving resilience against a pandemic: A novel technology for strategy development with practitioners and decision-makers. *Proceedings of the 20th International Conference on Information Systems for Crisis Response and Management* , Omaha, NE, USA.

Ackermann, F., Eden, C., Williams, T., & Howick, S. (2007). Systemic risk assessment: a case study. *J. Oper. Res. Soc.*, 58(1), 39-51.

Ackermann, F., Howick, S., Eden, C., & Williams, T. (2011). Delay and Disruption in Complex Projects. In R. A. Meyers (Ed.), *Complex Systems in Finance and Econometrics* (pp. 116-135), Springer.

Boaz, A., & Hayden., C. (2002). Pro-active evaluators – Enabling research to be useful, usable and used. *Evaluation*, 8(4), 440-453.

Bryson, J. M., Barberg, B., Carroll, A., C., E., George, B., Gonzalez, J. J., Rochester, J., Vandersmissen, L., and Zaki, B. (2023). Addressing Complex and Cross-Boundary Challenges in Government: The Value of Strategy Mapping. *The IBM Center for the Business of Government, Washington, DC, USA*. Retrieved February 5, 2024, from https://www.businessofgovernment.org/report/addressing-complex-and-cross-boundary-challenges-government-value-strategy-mapping.

Centeno, M. A., Nag, M., Patterson, T. S., Shaver, A., & Windawi, A. J. (2015). The Emergence of Global Systemic Risk. *Annual Review of Sociology*, 41(1), 65-85. doi:10.1146/annurev-soc-073014-112317.

Eden, C. Analyzing Cognitive Maps to Help Structure Issues or Problems. (2004). *European Journal of Operational Research*, 159(3), 673-686

Eden, C., & Gonzalez, J. J. (2023). The strategic management of disaster risk mitigation. In J. Radianti, T. Gjøsæter, & Y. Murayama (Eds.), *Information Technology in Disaster Risk Reduction ITDRR 2022* , (Vol. 672), Springer Nature.

Flood, S., Columbie, Y. J., Le Tissier, M., & O'Dwyer, B. (2022). Can the Sendai Framework, the Paris Agreement, and Agenda 2030 Provide a Path Towards Societal Resilience? In S. Flood, Y. J. Columbie, M. Le Tissier, & B. O'Dwyer (Eds.), *Creating resilient futures: integrating disaster risk reduction, sustainable development goals and climate change adaptation agendas*, Springer.

Gonzalez, J. J., Eden, C., Abildsnes, E., Hauge, M., Trentin., M., Ragazzoni, L., Berggren, P., Jonson, C.-O., and Abdelgawad, A. A. (2021). Elicitation, analysis and mitigation of systemic pandemic risks. *Proceedings of the 18th International Conference on Information Systems for Crisis Response and Management* , Blacksburg, VA, USA.

Gonzalez, J. J., & Eden, C. (2022). Insights from the COVID-19 Pandemic for Systemic Risk Assessment and Management. In J. Sasaki, Y. Murayama, D. Velev, & P. Zlateva (Eds.), *Information Technology in Disaster Risk Reduction ITDRR2021*, (Vol. 638), Springer Nature.

Gonzalez, J. J., & Eden, C. (2023). Devising Mitigation Strategies With Stakeholders Against Systemic Risks in a Pandemic. *Proceedings of the 20th International Conference on Information Systems for Crisis Response and Management*, Omaha, NE, USA.

Kaufman, G. G., & Scott, K. E. (2003). What Is Systemic Risk, and Do Bank Regulators Retard or Contribute to It? *Independent Review*, 7(3), 371-393.

Kim, Y.-k., Yoon, W. C., Lee, J., Poncelet, J.-L., Dolcemascolo, G., & Sohn, H.-G. (2022). A strategic response map for cascading pandemics: Lessons learned from the response to COVID-19 in the Republic of Korea. *Progress in Disaster Science*, 13, 100214. doi:https://doi.org/10.1016/j.pdisas.2022.100214

Lupton, D. (2013). *Risk* (2nd ed.), Routledge.

The Norwegian Corona Commision. (2020). *The authorities´ handling of the COVID-19 pandemic*. Retrieved 12th December 2023, https://www.regjeringen.no/contentassets/5d388acc92064389b2a4e1a449c5865e/no/sved/01kap02engelsk.pdf.

Norwegian Directorate for Civil Protection. (2020). *Analysis of Crisis Scenarios 2019*. Retrieved November 25, 2023, from https://www.dsb.no/rapporter-og-evalueringer/analyses-of-crisis-scenarios-2019/.

OECD (2018), *National Risk Assessments: A Cross Country Perspective*. OECD Publishing. https://doi.org/10.1787/9789264287532-en.

Oliver, K., Hopkins, A., Boaz, A., Guillot-Wright, S., & Cairney, P. (2022). What works to promote research-policy engagement? *Evidence & Policy*, 18(4), 691-713. doi:10.1332/174426421X16420918447616.

Reichstein, M., Riede, F., & Frank, D. (2021). More floods, fires and cyclones - plan for domino effects on sustainability goals. *Nature*, 592(7854), 347-349. doi:10.1038/d41586-021-00927-x.

Renn, O., Lucas, K., Haas, A., & Jaeger, C. (2019). Things are different today: the challenge of global systemic risks. *Journal of risk research*, 22(4), 401-415. doi:10.1080/13669877.2017.1409252

Schwarcz, S. L. (2008). Systemic risk. *The Georgetown Law Journal*, 97, 193-249.

Schweizer, P.-J. (2021). Systemic risks - concepts and challenges for risk governance. *Journal of risk research*, 24(1), 78-93. doi:10.1080/13669877.2019.1687574 .

Sterman, J. D. (2000). *Business dynamics: systems thinking and modeling for a complex world* , Irwin McGraw-Hill.

UNDRR & UNU-EHS. (2022). *Understanding and managing cascading and systemic risks: lessons from COVID-19*. Retrieved Dec. 15, 2023, https://www.undrr.org/publication/understanding-and-managing-cascading-and-systemic-risks-lessons-covid-19.

United Nations Office for Disaster Risk Reduction UNDRR. (2019). *Global Assessment Report on Disaster Risk Reduction 2019 (GAR2019)*. Retrieved December 5, 2023 from https://www.undrr.org/publication/global-assessment-report-disaster-risk-reduction-2019.

United Nations Office for Disaster Risk Reduction UNDRR. (2022). *Global Assessment Report on Disaster Risk Reduction 2022 (GAR2022)*. Retrieved December 5, 2023 from https://www.undrr.org/gar2022-our-world-risk.

Williams, T. M., Ackermann, F., & Eden, C. (1997). Project Risk: systemicity, cause mapping and scenario approach. In K. Kahkonen & K. A. Artto (Eds.), *Managing Risks in Projects* (pp. 343-352), E&FN Spon.

Perceiving and Behaving in a Crisis: Developing a Multi-Functional Crisis Information Platform for Psychosocial Situations (CIP-PS)

Sebastian Sterl
Technische Universitaet Braunschweig
sebastian.sterl@tu-braunschweig.de

Andreas Billig
Fraunhofer Institute for Open Communication Systems
andreas.billig@fokus.fraunhofer.de

Franck William Taffo
Fraunhofer Institute for Open Communication Systems
franck.william.taffo@fokus.fraunhofer.de

David Svoboda
Fraunhofer Institute for Open Communication Systems
david.svoboda@fokus.fraunhofer.de

Lars Gerhold
Technische Universitaet Braunschweig
lars.gerhold@tu-braunschweig.de

ABSTRACT

Increasingly long-lasting and complex crises and disasters can cause psychosocial effects on certain population groups with different vulnerabilities. To increase a population's resilience, it is necessary to examine these effects more deeply by systematically monitoring the psychosocial situation, integrating different types of data and making this data researchable, visible, and documentable via a multi-modal platform. Following a systems thinking approach, this paper presents a full approach to developing a "Multi-Functional Crisis Information Platform for Psychosocial Situations" (CIP-PS) for Germany, which includes the integration of an ontology- and filter-based research component (CIP-REP), a filter-based dashboard component (CIP-DAB), and a documentation component (CIP-DOC) into a complete information platform. In sum, this platform fulfills the requirement of comprehensively integrating different psychosocial data (e.g., survey, meta data, or question item data) into a system capable of fully linking all components and related data to healthcare informatics.

Keywords

Crisis information platform, psychosocial situation, data integration and visualization, dashboard and research platform, vulnerabilities

INTRODUCTION

The number of crises and disasters – such as, e.g., the Corona pandemic or extreme weather events – are increasing worldwide (Boin, 2019), posing risks and threats to the population (e.g., Munich Security Conference, 2023). When crises become more long-lasting, complex, intertwined, or transboundary (Boin, 2019), negative *psychosocial consequences* can emerge, such as post-traumatic stress disorder (PTSD), anxiety, or depression related to crises or disasters, which differ in relation to the vulnerabilities (*risk factors*) in certain population groups (Bonanno et al., 2010). This means that not only the physical condition of people affected by crisis or disaster is important when monitoring their effects. In coronavirus crisis management, for instance, psychosocial experts address the psychosocial situation of the population in addition to virological concerns (Hering et al., 2021), so as to increase the population's (psychosocial) resilience in times of crisis and disaster.

The main goal is thus to increase the population's resilience. Resilience is defined as a system's or an individual's ability to manage changes and continue developing (Moberg & Simonsen, 2011). The rise in complex and systemic crises leads us to take the approach of systems thinking: we understand society as a complex adaptive system (Carvalhaes et al., 2021) in which every stakeholder contributes to crisis preparation. This holistic approach is called the "system of systems" thinking (Cavallo, 2014). Additionally, Carvalhaes et al. (2021) speaks of a "wicked complexity" causing deep uncertainty and requiring new approaches to understanding how society functions. In order to assess resilience, combining quantitative and qualitative methods are seen as an innovative approach (Carvalhaes et al., 2021). According to the disaster resilience approach by Cavallo (2014), it is important to create an overall perspective, including top-down and bottom-up processes for unknown risks, and to regard crises management as an all-hazard-approach (Carvalhaes et al., 2021) to widen the situation picture, detect signals and see the bigger picture which is important to cope with the crisis (Leidner et al., 2009). Also, a shared understanding of the situation among all stakeholders as well as a modularization for software reuse is important for coordinating response activities and making it possible to reuse software in future projects (Leidner et al., 2009). This requires both a common knowledge base of psychosocial terms and situation definitions and a platform using individual components that can function alone and be applied in other crisis contexts. Furthermore, given the increasing importance of knowledge transfer, scientific knowledge should be prepared for target groups and co-developed in common projects aimed at actionable knowledge tailored to the needs of particular groups (Treffeisen & Grosfeld, 2022). Standards of transparency and verifiability ensure that society can trust this knowledge, which can be distributed on service platforms for the stakeholders concerned (Treffeisen & Grosfeld, 2022).

This raises the main research question:

> *How can an information platform contribute to increasing the population's (psychosocial) resilience in times of crisis and disaster?*

When appealing to the ability and resilience of citizens as their own crisis managers (see Boin, 2019), *individual resources* as well as the *perception* of and *behavior* in different crises should be monitored to derive suitable risk or crisis communication strategies to cope a crisis[1]. Also, other psychosocial information sources such as studies on help-seeking behavior in crises or field protocols on psychosocial emergency care have to be integrated into a system that monitors the overall psychosocial situation of the population at large as well as of those directly affected. Therefore, these different data concerning psychosocial topics should be combined in an information tool that makes them systematically researchable, visible, and documentable. That is, survey-based psychosocial situation pictures of the population as well as information on help-seeking behavior and psychosocial emergency care in times of crisis and disaster must be harmonized and fed into a platform targeting crisis managers as well as (government) decision-makers and authorities, the scientific and media community, and the public at large. This platform should be able to handle different psychosocial topics and related data accessed in different components depending on what the user is interested in. It must allow for an efficient and quick access to the components requested.

[1] We define the psychosocial situation picture using five psychosocial categories: perception (e.g., risk perception), behavior (e.g., protection behavior), risk factors (e.g., age) as well as resources (e.g., resilience) and psychosocial consequences (e.g., stress) (Gerhold et al., 2022; Lüttschwager et al., 2022).

This raises the more technical question:

> *How should an information platform be technically constructed so as to integrate different data types and connect different components with one another?*

Information, research, or dashboard platforms that collect and present various data in crises, disasters, or more general contexts do exist. While entire information platforms such as INFORM (European Commission - Disaster Risk Management Knowledge Centre [EC-DRMKC], 2023) or dashboards such as the Natural Disasters Dashboard (Office for the Coordination of Humanitarian Affairs [OCHA], 2023) use statistical or process data, only few information or dashboard platforms use survey data like the Smart Data Disaster Management (Smart Data, 2023) or PACE (University of Erfurt et al., 2023). Research platforms with functionalities similar to ours are rare. They include the IFRC GO (Red Cross Red Crescent [IFRC], 2023) for disasters in general or the paneldata.org-explorer (German Socio-economic Panel [GSOEP], 2023) in generic contexts using statistical/process or survey data, among others.

To the best of our knowledge, there are to date no information platforms combining survey, systematic review, and psychosocial emergency care data.

In terms of application, the question becomes:

> *How can different data be integrated and aggregated into a tool that researches, visualizes, and documents them?*

We answer these research questions by first explaining our research methodology, then reviewing the current research, including an overview of existing information, research, and dashboard platforms, before showing in greater detail how we developed the "Multi-Functional Crisis Information Platform for Psychosocial Situations" (CIP-PS). Therefore, we explain the overall structure and outline the data used for integration into the entire system. Next, we describe the single elements of the research component CIP-REP, the dashboard component CIP-DAB, and the documentation component CIP-DOC, including their technical and visual implementation and how they link to one another. We then present different use cases for CIP-REP and CIP-DAB in different scenarios. The final sections discuss challenges and limitations, followed by a conclusion and future steps.

RESEARCH METHODOLOGY

In this section, we explain our research methodology, following those in information systems (IS) summarized in Kilani and Kobziev (2016) and the design science research[2] IS framework by Peffers et al. (2007). First, we defined the problem as a lack of information platforms providing information on the psychosocial situation in crises using state-of-the-art research. We addressed this in several research steps, such as tool programming and data collection/processing, development, and prototyping (Archer, 1984). In general, we followed a two-stage approach, by looking at CIP-PS as a whole (stage 1) and then its components CIP-REP, CIP-DAB, and CIP-DOC (stage 2), using both a qualitative and a quantitative approach.

In stage 1, we exploratively developed CIP-PS by means of data collection steps such as state-of-the-art research on established information platforms (see below), and a target group analysis (with our project partner from the German Red Cross as the first target group). Next, we created an overview of the psychosocial topics and conducted further group interviews with aid organization members in order to explore the potential use of our platform. Our research strategy entailed multiple use cases (shown below), addressing precise potential problems, target groups, and processes. The (technical) validity and reliability of the platform was ensured by applying scientific standards and established solutions of User Interface (UI) and User Experience (UX) research in its development. Components from other projects were reused to create synergy effects. CIP-PS systematically presents psychosocial topics in order to illustrate the complex psychosocial crisis situation and reduce complexity in a model that can be analyzed and that explains the situation and its interdependencies in an understandable and transparent way. This serves target groups as a decision support system and improves individual resilience, the core purpose of CIP-PS (as set forth in the Introduction).

In stage 2, the different components as a "system of systems" (see above) are treated with different sub-methodologies.

CIP-REP follows a qualitative approach by integrating results of documents and study outcomes using meta-

[2] Simon (1969, p. 55) as cited in Peffers et al. (2007) stated, that "[w]hereas natural sciences and social sciences try to understand reality, design science attempts to create things that serve human purposes".

analysis and the qualitative content analysis research strategy (Mayring, 2015). The terms of these documents were deductively derived into ontologies, filters, and contents and can be inductively expanded with new concepts.

CIP-DAB, by contrast, follows a quantitative research strategy using quantitative data on surveys and protocols, which are deduced from the framework of psychosocial situation research (e.g., Gerhold et al., 2022) and represented in validated and reliable question constructs. These are presented in a statistically reproducible and clear way (e.g., plots) differentiated by scientifically derived risk factors.

Finally, CIP-DOC applies scientific principles of transparency and documentability to both components by storing easily accessible documentation of studies, questionnaires, and protocols.

STATE OF THE ART AND RELATED WORK

We define an information platform as a "software-based platform" and core functionality with different modules as software subsystems which can then "extend the functionality of the platform" (Hein et al., 2020, p. 88). The information platform as an overall framework embeds sub-systems such as CIP-REP as well as CIP-DAB. CIP-REP is an IR (Information Retrieval)-based component for semantics-enabled searches of project artifacts, whereas a dashboard is a visualization "of a consolidated set of data for a certain purpose" (Matheus et al., 2020, p. 1) with different types of presentations, such as geographic, numerical, or temporal (Schulze et al., 2023).

Table 1 illustrates some examples of information platforms like the INFORM (EC-DRMKC, 2023), HEIMDALL (German Aerospace Center [DLR], 2021), or the Global Crisis Response Platform (International Organization for Migration [IOM], 2023) that integrate different elements and crisis-related data into one platform. Research platforms such as the IFRC GO (IFRC, 2023) or Google Crisis Response (Google, 2023) allow users to systematically research the data that is also presented visually, while general research platforms such as paneldata.org (GSOEP, 2023) enable the user to search for survey variables similar to our CIP-REP component with its psychosocial situation monitoring survey data. In terms of dashboards, there are examples of tools integrating survey data, such as the PACE project (University of Erfurt et al., 2023) or the Covid-19 National Survey Dashboard (CAMH, 2023), with many others visualizing statistical or process data on the environment, natural disasters, or crises in general (e.g., Community Risk Assessment: Netherlands Red Cross, 2023; Natural Disasters Dashboard: OCHA, 2023; Global Crisis Risk Dashboards: United Nations Development Programme, 2020). Another selection of dashboards can be found in Sterl et al. (2023) or, for public health data, in Schulze et al. (2023).

Table 1. Examples of information platforms (MIP), research platforms (REP), and dashboards (DAB)

Type of platform	Type of crisis	Type of data	Specific content	Example of tool (Citation)
MIP	Environment/climate/humanitarian crises	Statistical/process	Information about crises in general (focus on, e.g., risks, warnings, climate change)	INFORM (EC-DRMKC, 2023)
	Natural disasters	Statistical/process	Disaster-relevant information	HEIMDALL (DLR, 2021)
	Medical crises	Statistical/process	Information about coordination and decision processes (concerning crises/medical emergencies)	COncORDE – Crisis Management Platform (Konnektable Technologies, 2021)
	General disasters	Statistical/process/survey	Real-time data on crisis-relevant information (e.g., water levels, stress levels, social	sd-kama (Smart Data, 2023)

			network structure)	
	General crises	Statistical/process	Crises-relevant information	Global Crisis Response Platform (IOM, 2023)
REP/DAB	Disasters in general	Statistical/process	Disaster-relevant information	IFRC GO (IFRC, 2023)
	General crises	Statistical/process	Crises-relevant information	Google Crisis Response (Google, 2023)
REP	General	Statistical/process/survey	Information on, e.g., research data	paneldata.org (GSOEP, 2023)
	General	Statistical/process/survey	Information on, e.g., research data	GESIS search (GESIS - Leibniz Institute for the Social Sciences, 2023)
DAB	Covid-19	Survey	Covid-19 information (main focus on the mental health and substance use in Canadian population)	COVID-19 National Survey Dashboard (CAMH, 2023)
	Environment/climate crisis	Survey	Information about the climate crisis (focus on, e.g., attitudes, behavior, risk perception)	PACE Explorer (University of Erfurt et al., 2023)
	Environment/climate/ humanitarian crises	Statistical/process	Information about affected areas in humanitarian crises	Community Risk Assessment (Netherlands Red Cross, 2023)
	Natural disasters	Statistical/process	Disaster-relevant information	Natural Disasters Dashboard (OCHA, 2023)
	Ukraine crisis	Survey	Information about the Ukraine crisis	GeoPoll Ukraine Crisis Data Tracker (GeoPoll, 2022)
	General crises	Statistical/process	Information on global risks	Global Crisis Risk Dashboards (United Nations Development Programme, 2020)
	General crises	Statistical/process	Crises-relevant information (e.g., geographical,	The Emergency Dashboard (World Food Programme,

	funding, and performance-related information)	2023)

Note: This represents only a small sample. Another selection can be found in Sterl et al. (2023); a much wider selection can be requested from the authors of this paper.

In sum, there are many platforms and components that focus on crises, disasters, or general events. However, most tools primarily integrate process-produced information instead of population surveys, literature reviews of scientific studies, or information on psychosocial emergency care that would present an overall picture of the psychosocial situation in times of crisis and disaster.

To explore the information items concerning survey and statistical data, mature techniques can provide adequate means for a broad range of target users.

The approaches presented here focus on the integration of three diverse, almost independent views (see Sterl et al., 2023):

1) Research and exploration offering an interface for the semantic retrieval of domain-relevant data
2) Statistical interdependencies with CIP-DAB utilizing enhanced visualization techniques for statistical data
3) Documentation library, inspecting static, document-oriented artifacts

Given the above, we see two gaps in the research:

First, based on the systems thinking approach and the knowledge transfer paradigm, we bring together different data on psychosocial topics into one platform to reduce complexity, and to help users systematically research, visualize, and document all the data in a way that is tailored to their needs and addresses the main goal of increasing resilience in the population.

Next, our IS fully connects not only a dashboard visualization component but also a research and documentation component in the front as well as backend layer to help the users visualize the information required for their decisions.

We conclude that the development of CIP-PS is – to the best of our knowledge – innovative and fills in the research gaps we noted above. The next sections present in greater detail the overall structure of CIP-PS as well as its single components CIP-REP, CIP-DAB, and CIP-DOC.

DEVELOPING THE CRISIS INFORMATION PLATFORM CIP-PS

General Concept

CIP-PS is divided into three different layers, including the psychosocial topics: “Psychosocial situation in Germany” (PSG), “Help-seeking behavior” (HSB), and “Psychosocial emergency care” (PEC) (*Figure 1*). Depending on what topic the user selects, the respective CIP-REP, CIP-DAP, or CIP-DOC opens. For both PSG and PEC, CIP-REP, CIP-DAB, and CIP-DOC are displayed, while for HSB, CIP-REP and CIP-DOC are shown. Further, descriptions of each psychosocial topic are shown after clicking on the appropriate information text section.

The overall platform design is structured as a micro frontend. That is, CIP-REP, CIP-DAB, and CIP-DOC are micro frontends independent of each other and linked via an integration layer. The complete frontend is composed of different, independently working components. The data which is called up by the user is pre-processed using a proxy server and then integrated into a data format representing the specific characteristics of every frontend. CIP-REP, CIP-DAB, and CIP-DOC can be deployed singularly regarding the other elements, while CIP-PS is highly responsive and restricted to tablet mode following the principles of high UI/UX standards using the bootstrap 4 library with their own guidelines (Bootstrap, 2023). Regarding app functionality, the CIP-PS is designed as a Progressive Web App (PWA; Tandel & Jamadar, 2018) working offline. This means that it loads all the required data when it is first established or called up by the user, and thus runs much more quickly in subsequent single uses. The data required in the respective component is saved in the background so that the user can keep researching free from interference. Since a PWA also works offline running directly on the device, the CIP-PS can be used without an Internet connection.

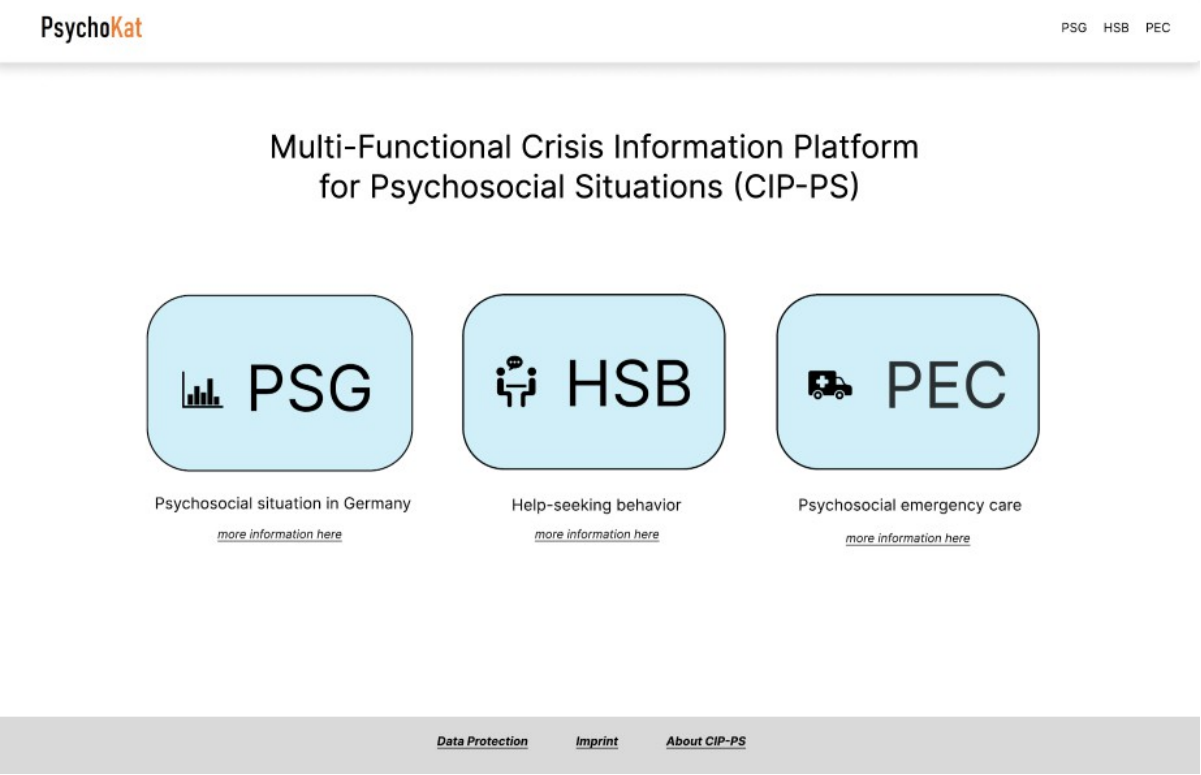


Figure 1. Screenshot of the overall structure of CIP-PS

Research Component: CIP-REP

CIP-REP is dynamic and includes filter- as well as ontology-based research on PSG, HSB, and PEC (see *Table 2* and *Figure 2* below). For PSG, data includes questionnaires, codebooks, as well as descriptive statistics of both the national and model regions per survey wave. Users can search psychosocial variables using the search field or a topic or word cloud. As a result, for instance, variable names, questions, and value labels as well as descriptive univariate statistics are shown in the form of an overview. For PEC, data consists of documentation (e.g., field/operation protocols, research, literature). For HSB, there are scientific studies on crises and disasters. Certain words or phrases can be used to find PEC or HSB results filtered by, e.g., year of publication or research method. Data must be uploaded manually and can be expanded with further information in the form of, e.g., questionnaires for PEG.

Table 2. Structure of CIP-REP

Psychosocial topic	1. PSG 2. PEC 3. HSB
Integrated data by psychosocial topic	1. Codebooks, questionnaires and descriptive calculations of national/model region surveys per survey wave 2. Documentation on PEC (e.g., field/operation protocols, research, literature) 3. Documentation on HSB in crises and disasters (e.g., individual studies)
Result displays (e.g.)	1. Question wording, question values, type of crisis 2. Publication year, title, author 3. Publication year, research method, population group
Overall content use/data upload/data base	- Dynamic/manual/expandable
Searches	- Filter- and ontology-based
Additional remarks	- Ontology representation according to SKOS standard, additional authoring component

One of the technical requirements for the multi-modality of CIP-REP is a component design which should be particularly integrative. For the core information model of CIP-REP based on the integrated data in *Table 2* (e.g., codebooks, questionnaires, or descriptive calculations), it is important to avoid duplicating shared information structures and increase and maximize the reprocessing of common structures, especially patterns. Apart from these classic data-oriented information structures, ontologies as well as the corresponding search and exploration of the underlying data are crucial attributes. Various domain ontologies, based on different data sources and represented by the SKOS standard (World Wide Web Consortium [W3C], 2012), stand for the specific vocabularies of the particular domain. The multi-ontology approach focuses on genericity in order to dynamically support exchanging domain ontologies. This can be separately applied to various domains, e.g., the medical domain (Billig & Krebs, 2014). Here we adapted mature partition techniques from Decker et al. (2005) and ontology matching techniques from Billig et al. (2007).

Dashboard Component: CIP-DAB

CIP-DAB is also structured in a dynamic and interactive way and uses filters for the psychosocial topics PSG and PEC (see *Table 3* and *Figure 3* below). For PSG, it involves longitudinal survey population data on the psychosocial situation, cross-sectional survey data on preparedness for special crises (e.g., heat, terrorist attack), and model region survey data on PEC per survey wave in Germany. The results include statistically pre-processed longitudinal data (e.g., CI-plots, bar charts) on psychosocial variables such as risk perception, protective behavior, or coping strategies collected for diverse crises (Ukraine crisis, Covid-19, climate change, or inflation) by risk factor (e.g., age or sex) or region (e.g., German states). For PEC, there is simulated data on PEC field protocols and information on PEC units of an aid organization operating in German communities, which then are displayed descriptively using univariate and bivariate statistics (e.g., bar charts or tables). Data such as survey data for PSG, for instance, must be uploaded manually but can be expanded with further information.

Table 3. Structure of CIP-DAB

Psychosocial topic	1. PSG 2. PEC
Integrated data by psychosocial topic	1. Longitudinal survey data on the psychosocial situation, cross-sectional survey data on preparedness for special crises (e.g., heat, terrorist attack), model region survey data on PEC 2. Simulation survey data on PEC field protocols, information on PEC units of aid organization in German communities
Result displays (e.g.)	1. Univariate/bivariate statistics on survey variables, risk factors, region (e.g., CI-plots, bar charts, summary) 2. Univariate/bivariate statistics on simulated variables of emergency care field protocols (e.g., CI-plots, bar charts, summary); descriptive information on PEC units
Overall content use/data upload/data base	- Dynamic/manual/expandable
Searches	- Filter-based
Additional remarks	- Contains a map, graphs, data input; uses libraries such as leaflet map, Bootstrap-select or D3.js

The dashboard consists of three main components using various libraries, for data input, map, and graph section. For the data input, five select elements were built with the Bootstrap library (Otto & Thornton, 2023). Further, using a map section, the affected regions can be displayed on a map with polygons of varying thickness. The map was implemented with Leaflet (Agafonkin, 2023) and the polygon data is provided by a GeoServer. Finally, the graphs, which serve as the main response view, were implemented separately with the D3.js library (Mike Bostock and Observable, 2023) and are added to the project via dependency injection. Some of the advantages of CIP-DAB are the ability to separately deploy the visual controls of the dashboard as well as its capacity for working with a large amount of data.

Documentation Component: CIP-DOC

CIP-DOC is a static tool for downloading information on PSG and PEC in order to quickly obtain required files (see *Table 4*). For PSG, survey information such as questionnaires, codebooks, and short descriptive reports of survey waves are integrated and downloadable in a readable and portable file format. For PEC, data comprises documentation on PEC (e.g., standardized field/operation protocols, research, literature) which can be downloaded as a ready-to-read file. Finally, all the data is uploaded manually and expandable by adding further information.

Table 4. Structure of CIP-DOC

Psychosocial topic	1. PSG 2. PEC
Integrated data by psychosocial topic	1. Codebooks, questionnaires, and short survey reports 2. Documentation on PEC (e.g., standardized field/operation protocols, research, literature)
Result displays (e.g.)	1. Readable and portable files of codebooks, questionnaires and short survey reports 2. Readable and portable files of documentation on PEC (e.g., field/operation protocols, research, literature)
Overall content use/data upload/data base	- Static/manual/expandable
Additional remarks	- Download function, different search methods

The documentation page is a pure, static HTML with a list of cards showing information with the help of the Bootstrap library (Otto & Thornton, 2023). Each card shows a summary of the information; for a detailed view, the user has to click on the card. The information on the documentation page is stored on the server and can be updated, with each card displaying its update date. Furthermore, users can type terms they are searching for into an input field, with a meta data scheme in the backend. Running in the background and not visible for the user, a local search in the meta data or a global search using the texts themselves is also possible. In addition, CIP-DOC contains filters differentiated by psychosocial topic. For CIP-REP, as explained above, CIP-DOC delivers textual artifacts for expanding the domain ontologies used in the CIP-REP component.

USE CASES

CIP-PS addresses crisis managers, such as decision-makers from aid organizations, public authorities (e.g., the Red Cross or public health institutes), and government (small regions to federal states). Thus, CIP-PS is a decision support system for those directly working with the psychosocial situation in times of crisis and disaster. Furthermore, representatives from the media and academia as well as citizens at large can access CIP-PS for a descriptive overview of the situation. The following use cases show different applications of CIP-PS.

CIP-REP: Researching Variables and Concepts

In use case 1, the target group includes a researcher (RE) in the psychosocial field wanting to use our survey data for scientific purposes. These secondary data are collected and stored under the highest data protection and anonymity standards. RE is interested in the research question "Which factors are relevant for private emergency preparedness in a heat event?" and is guided by the protection motivation theory (Rogers & Prentice-Dunn, 1997) in which risk perception is key. As RE does not know what variables are in our dataset, they choose "risk perception" as a construct in CIP-REP (see *Figure 2*). If this variable exists, RE would like to know, what *category* it is in, the *wording of the question*, *values*, the *type of crisis*, and the *survey waves* the variables were measured for. According to *Figure 2*, RE's search can be narrowed to specific topics shown on the left. Moreover, the word cloud in the center as well as the filters on the left include specific characteristics to further narrow down the search, thereby bringing about results more effectively and efficiently. Upon completion of the search, the results display all the variables in our dataset matching the criteria RE has provided. Now, RE has a list of possible variables to analyze and could use our data for this purpose.

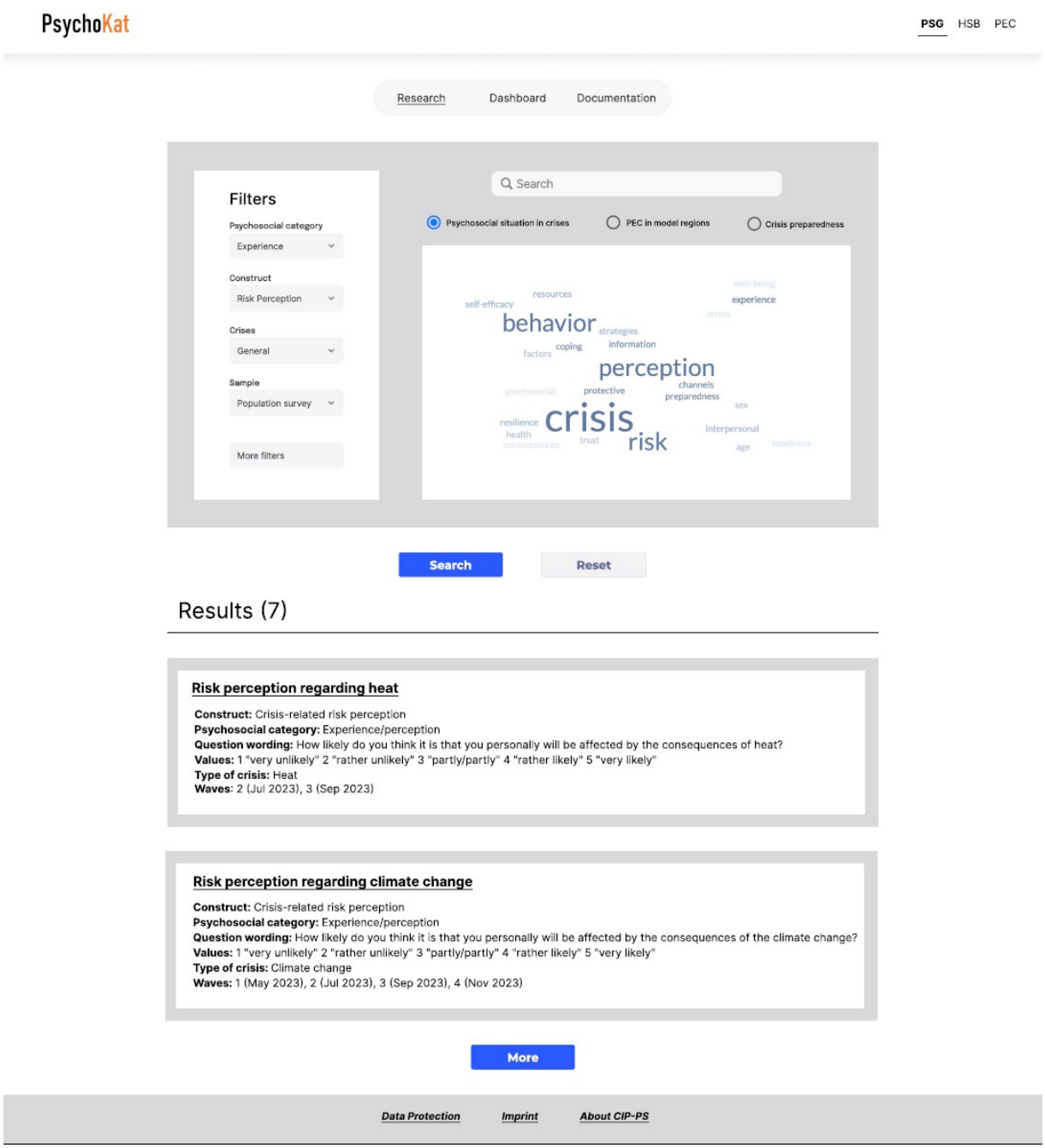


Figure 2. Screenshot of CIP-REP (use case 1)

CIP-DAB: Visualizing Risk Perception and PEC in Multiple Crises

In use case 2, the user is a crisis manager (CM) in an authority. CM wants to develop an effective risk communication strategy tailored to specific at-risk groups and is interested in the current state of the population so to be able to identify vulnerable groups (Bundesamt für Bevölkerungsschutz und Katastrophenhilfe [BBK] & Bundesinstitut für Risikobewertung [BfR], 2022). To this end, CM searches for information shown on CIP-DAB, focusing on the risk perception of different population groups, differentiated geographically as well as historically. As shown in *Figure 3*, CM can select "Crisis-related risk perception" as a psychosocial variable to monitor. Also, CM is interested in age group as a risk factor in all German states. Furthermore, CM selects climate change out of all crises measured since May 2023 as a trend survey design. *Figure 3* shows the input fields, as well as a map with different values of risk perception by age group using a mouseover effect as well as a CI-plot and bar plot as visual controls (among other visuals such as a summary table of statistical values and a correlation matrix). Using the CI-plot, CM can tell if there are significant differences in mean values of risk perception between different age groups at one survey time and between different survey waves. Further, the bar plot includes absolute as well as relative frequencies of risk perception (using a mouseover effect) differentiated by age group and survey wave.

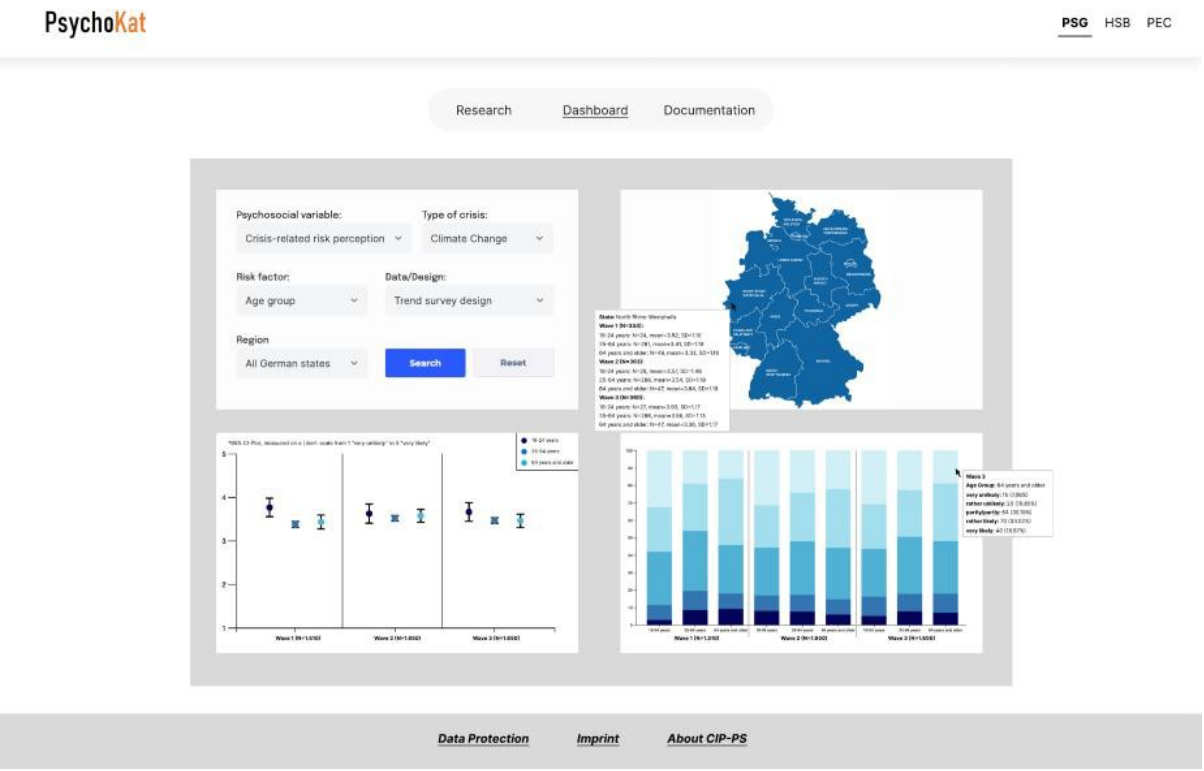


Figure 3. Screenshot of CIP-DAB (use case 2)

For the sake of clarity, we omit a figure for use case 3. Here, the head of the community level (HC) of PEC is interested in survey data on the knowledge about and personal usage of PEC in crises. The goal of HC is to advertise the service of PEC in these regions and to determine whether more resources are needed within the aid organization's structure. This information was measured for several model regions where PEC is offered by the aid organization. Thus, HC selects model region survey data in CIP-DAB obtaining the questions "Before participating in this survey, did you know what PEC is?" and "Have you ever used PEC from the respective aid organization in your region?" which will immediately provide HC with information that will help ground decisions on the investment of further resources.

CHALLENGES AND LIMITATIONS

Challenges

A general challenge is to accordingly align the technological development of the system components to the reuse of information substructures; namely, because joint patterns within the general information structure can be identified. For example, different ontologies, stemming from the three psychosocial topics (see *Figure 1* and *Table 2*), can share common concept spaces. Furthermore, common metadata sub-structures have to be identified and integrated accordingly.

One challenge of developing the dashboard is designing it as a micro frontend that has not been established yet. Its components were implemented using various libraries as shown above and function independently of each other. The great advantage of this micro-frontend architecture is that the dashboard can be deployed outside CIP-PS. The diagrams can also be deployed individually as a widget on a platform. This means that, for instance, the map would not cause a complete failure of the dashboard as happens with standard websites. The number of output diagrams can be expanded, and the individual diagrams can be built with different libraries.

A further design challenge is posed by the structuring and designing of the single elements within CIP-PS. The aim is for all elements to be customized to the needs of the user according to the highest standards of UI and UX. The platform should be easily understandable, handy, and user-friendly, so that the user (e.g., crisis managers, authorities, decision-makers) can obtain all the information they need and use it as a support system for their evidence-based decisions. At the same time, the principle of parsimony should ensure that using CIP-PS should not lead to information overload. For instance, an onboarding process with help texts would support this aim.

As concerns privacy regulations, the platform follows the principle of different user roles, meaning that it can be secured by using a safe login procedure. Further, all the data is stored on secure servers and is not available without permission.

As concerns data challenges, the data which is sent to CIP-PS and its individual components must be pre-processed. Using a data-centered approach, the national and model region survey PSG data is prepared and saved into single data files with a high number of possible combinations (e.g., different psychosocial variables by different risk factors). Data on PEC and HSB comprises studies and other documentation that have to be reviewed manually before saving all the information required for the different filters into differently structured tables.

The growth of "computing power, more complex algorithms and the availability of ever more potentially complimentary data online" (Quinn & Malgieri, 2021, p. 1583) means that the amount of sensitive data is increasing. This leads to social challenges, such as the need to protect sensitive data to prevent risks such as discrimination and harm to vulnerable groups in society. Naturally, there are high data protection standards when it comes to collecting survey data. This type of data is then also handled carefully, is anonymized and aggregated to make sure it cannot be traced to individual respondents, and is stored on a secure server. For the secondary use of the survey data, regional information such as postal codes or other identifying information about respondents is left out to ensure total anonymity, as a combination of only a few characteristics can potentially lead to identification (Bampoulidis & Lupu, 2019). For HSB, the individual underlying studies are not provided in the CIP-REP, but are freely accessible under scientific license conditions. The PEC literature is also freely accessible and no sensitive data, such as interview protocols, are stored in the CIP-REP or CIP-DOC. The PEC data for CIP-DAB consists of completely simulated data on protocols (no real data) in order to protect privacy, while the survey model region data is also aggregated and secured.

Limitations

Naturally, the development of CIP-PS is also subject to some limitations. Next, this section concerns limitations regarding the data of the psychosocial topics and general limitations in the technical development of CIP-PS and CIP-REP, CIP-DAB and CIP-DOC.

In terms of PSG, we have collected data in short intervals online-representative of sex, age, and German states, both as concerns the national situation as well as the three model regions. Although collected quite frequently, this data is nevertheless limited in time and space when it is integrated into CIP-DAB. Furthermore, the model regions were selected according to criteria on the occurrence of PEC teams, population size, and the number of single districts we can then apply multi-level models on. But these regions are not representative of all locations in Germany. Thus, we cannot monitor the psychosocial situation in (rural) regions, communities, or smaller

locations where disasters such as extreme weather events or wildfires (e.g., the flooding in Germany in 2021 or wildfires in 2023) occur. Moreover, CIP-DAB data only represents the German population, while disaster and crisis events happen all over the world, having more or less of an impact on the psychosocial situation of the population affected or involved. It would thus be essential to develop a cross-cultural international research network studying different perceptions and behaviors of the various populations in different or same crisis or disaster events.

In terms of PEC, only simulation data of operation protocols can be generated and integrated into CIP-DAB for data protection and security reasons. At this stage, it is not possible to process data from real people affected or involved in crisis or disaster situations. We will respond to this by developing an interface facilitating the connection between real data and CIP-DAB, with the aim of creating a framework with a full data protection scheme.

One limitation with respect to the technical development is that the data used by the components CIP-REP and CIP-DOC cannot be dynamically changed during system runtime. Thus, for example, artifacts such as questionnaires, studies, etc. have to be maintained offline. Furthermore, there are limitations in the development of the domain ontologies: first, the language expressivity of ontologies for CIP-REP is restricted to SKOS expressivity; second, new versions of ontologies cannot be dynamically added at runtime.

Although CIP-DAB is relatively flexible, there is a rather passive limitation when it comes to the resolution. That is, CIP-DAB is responsive in tablet mode only and is not necessarily suitable for mobile phones. Furthermore, CIP-PS and CIP-DAB both need an Internet connection when updating results. The PWA, however, as described above, helps keep recent searches stable before updating.

CONCLUSION AND OUTLOOK

In this paper, we described in detail the full development of our CIP-PS. In a first step, we introduced our conceptual foundation by explaining why it is important to monitor the psychosocial situation of the population during times of crisis and disaster, our systems thinking approach to increasing individuals' resilience, and which psychosocial topics need to be addressed. Thus, psychosocial aspects such as monitoring the situation to identify vulnerabilities in the population as well as summarizing studies on HSB and PEC information of aid organizations are crucial for drawing an overall picture of the situation. After explaining our two-stage research methodology approach, we presented a list of relevant recent platforms (see *Table 1*). This made clear the need for an information platform that follows a highly integrative approach.

Our main research question was: *"How can an information platform contribute to increasing the population's (psychosocial) resilience in times of crisis and disaster?"* The platform as IS delivers information to find correlations in different psychosocial data through an integration of different components. As both crises and the psychosocial situation are becoming more complex, the platform looks at statistical relationships in the data, reduces complexity and builds a system of systems (Cavallo, 2014, our three components) to model (psychosocial) reality at a manageable level. By integrating different components into the system of CIP-PS, it is thus possible to reach potential user groups from crisis managers to researchers and the public, thus fulfilling a policy need as well as responding to a research interest. Furthermore, spill-over, crowding-out, or habituation effects (Martin-Lapoirie et al., 2024; Weber, 2010) in different crises mean that there is a need to compare the concurrent behavior and perception of the population in multiple crises represented in CIP-PS. Other IS on surveys look at one crisis only (e.g., PACE, University of Erfurt et al., 2023), so that information has to be collected from different sources when developing action plans or communication strategies for the population at risk. Moreover, integrating different components improves statistical dependencies between them: thus when, for example, PSG in CIP-DAB shows that the elderly are more vulnerable in a particular situation, CIP-REP can help identify facilitators for appropriate HSB to further increase or restore the vulnerable group's resilience. Studies in CIP-REP, in turn, can lead to a derivation of new psychosocial variables to monitor in PSG. In conclusion, the different levels of complexity of crises and disasters make it difficult to identify people or population groups who are more or less vulnerable and foster or maintain their resilience accordingly – hence the need to monitor psychosocial consequences in the long run using an integrative approach of a system of systems.

The second question: "*How should an information platform be technically constructed so as to integrate different data types and connect different components with one another?"* addresses a technical aspect of the development of CIP-PS. We answered it by explaining the general concept of CIP-PS, its design framework (see *Figure 1*), as well as the technical functionalities concerning micro frontends, including the use of a PWA. We also gave a technical description of CIP-REP, CIP-DAB, and CIP-DOC.

We answered the third research question: "*How can different data be integrated and aggregated into a tool that researches, visualizes, and documents them?"* by describing the psychosocial topics and respective data each individual component integrates as well as how they display results and provide ways to edit and add further data into the systems (see *Tables 2–4*). Also, the functionality of CIP-REP and CIP-DAB was illustrated by three different use cases differentiated by user group, scenario, and data used, in order to show the diverse situations, the multi-functional crisis platform can be used for (see *Figures 2* and *3*).

While we have presented here a full development of our CIP-PS, we also want to mention future steps. First, additional national and model region survey data on several subsequent waves and short survey reports will be collected, processed, and sent to CIP-PS in order to expand all three components. The ontologies based on the data on PSG, PEC, and HSB will also be expanded to make the search using CIP-REP more efficient and effective. In addition, we plan further publications with both a theoretical and methodological focus (e.g., developing further the ontologies used or making the platform more responsive) in order to address gaps in the research.

In technical terms, the multi-modal system can be extended with respect to CIP-REP and CIP-DOC. In particular, integrative aspects like interlinking of semantically similar artifacts of CIP-REP and CIP-DOC could play a prominent role. Furthermore, an authoring element for all three CIP components and their related data could be established. Beyond the features of a classic Content Management System (CMS), this authoring element would support cooperative editing of, for example, codebooks, studies, or ontologies tailored to their individual structures. Further UI and UX improvements are planned for CIP-DAB.

All in all, we conclude that our CIP-PS addresses two major research gaps: first, based on the systems thinking approach and the knowledge transfer paradigm, it links different data on psychosocial topics together into one platform, thus reducing complexity. It helps the users systematically research, visualize, and document all the data, which is customized in favor of their needs and requirements, with the main goal of increasing resilience in the population. Second, CIP-PS proofs a full connection of a CIP-DAB visualization as well as a CIP-REP and CIP-DOC in both the frontend and backend, helping the users to obtain and visualize the information required as a decision support system. Hence, it is essential to present psychosocial data by topic independently as well as efficiently and effectively in the sense of a tool valuable for disaster public health and healthcare informatics.

ACKNOWLEDGMENTS

We would like to thank our research intern Lea Malcharczyk for her invaluable support in writing this article.

REFERENCES

Agafonkin, V. (2023). *Leaflet — an open-source JavaScript library for interactive maps*. https://leafletjs.com/

Archer, L. B. (1984). Systematic method for designers. In N. Cross (Ed.), *Developments in Design Methodology*. (pp. 57–82). John Wiley.

Bampoulidis, A., & Lupu, M. (2019). *An Abstract View on the De-anonymization Process.* https://doi.org/10.48550/arXiv.1902.09897

Billig, A., Blomqvist, E., & Lin, F. (2007). Semantic Matching Based on Enterprise Ontologies. In R. Meersman & Z. Tari (Eds.), *Lecture Notes in Computer Science. On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS* (Vol. 4803, pp. 1161–1168). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-76848-7_76

Billig, A., & Krebs, F. (2014). *Ein wissensbasierter Terminologie-Dienst zur Unterstützung von Konzept-getriebenen E-Health-Prozessen*. https://doi.org/10.24406/publica-fhg-386563

Boin, A. (2019). The Transboundary Crisis: Why we are unprepared and the road ahead. *Journal of Contingencies and Crisis Management*, *27*(1), 94–99. https://doi.org/10.1111/1468-5973.12241

Bonanno, G. A., Brewin, C. R., Kaniasty, K., & La Greca, A. M. (2010). Weighing the Costs of Disaster: Consequences, Risks, and Resilience in Individuals, Families, and Communities. *Psychological Science in the Public Interest : A Journal of the American Psychological Society*, *11*(1), 1–49. https://doi.org/10.1177/1529100610387086

Bootstrap. (2023). *Bootstrap 4 - Ref. System on UI Guideline*. https://www.uiguideline.com/reference-systems/bootstrap-4

Bundesamt für Bevölkerungsschutz und Katastrophenhilfe, & Bundesinstitut für Risikobewertung. (2022). *Risikokommunikation [Risk Communication]: Ein Handbuch für die Praxis [A handbook for practical use]*. https://www.bbk.bund.de/SharedDocs/Downloads/DE/Mediathek/Publikationen/Risikomanagement/handbuch-risikokommunikation.pdf? __blob=publicationFile&v=3

CAMH. (2023). *COVID-19 National Survey Dashboard*. https://www.camh.ca/en/health-info/mental-health-and-covid-19/covid-19-national-survey

Carvalhaes, T. M., Chester, M. V., Reddy, A. T., & Allenby, B. R. (2021). An overview & synthesis of disaster resilience indices from a complexity perspective. *International Journal of Disaster Risk Reduction*, *57*, 102165. https://doi.org/10.1016/j.ijdrr.2021.102165

Cavallo, A. (2014). Integrating disaster preparedness and resilience : a complex approach using System of Systems. *The Australian Journal of Emergency Management*, *29*(3), 46–51. https://search.informit.org/doi/10.3316/agispt.20142407

Decker, S., Sintek, M., Billig, A., Henze, N., Dolog, P., Nejdl, W., Harth, A., Leicher, A., Busse, S., Suess, J. G., Miklos, Z., Ambite, J. L., Weathers, M., Neumann, G., & Zdun, U. (2005). *TRIPLE - an RDF Rule Language with Context and Use Cases: In W3C Technology and Society domain (Chair), W3C Workshop on Rule Languages for Interoperability, 27-28 April 2005, Washington, DC, USA.* http://www.w3.org/2004/12/rules-ws/paper/98

European Commission - Disaster Risk Management Knowledge Centre. (2023). *INFORM - Global, open-source risk assessment for humanitarian crises and disasters*. https://drmkc.jrc.ec.europa.eu/inform-index

GeoPoll. (2022). *GeoPoll Ukraine Crisis Data Tracker*. https://www.geopoll.com/resources/geopoll-ukraine-crisis-data-tracker/

Gerhold, L., Stelzmann, D., Sterl, S., & Lüttschwager, N. (2022). Psychosoziales Lagebild der Bevölkerung während der Corona-Pandemie (LaBeCo) : Abschlussbericht zum BMBF-Forschungsprojekt : BMBF-Fördermaßnahme: "Zivile Sicherheit - Forschungsansätze zur Aufarbeitung der Corona-Pandemie" : Laufzeit des Vorhabens: 01.01.2021-31.12.2021 [Psychosocial situation picture of the population during the Corona pandemic (LaBeCo) : Final report on the BMBF research project : BMBF funding measure: "Civil security - research approaches to dealing with the Corona pandemic" : Duration of the project: 01/01/2021-12/31/2021. Psychosocial situation of the population during the Corona pandemic (LaBeCo).]. *Psychosoziales Lagebild der Bevölkerung während der Corona-Pandemie (LaBeCo).* Advance online publication. https://doi.org/10.2314/KXP:1851570950

German Aerospace Center. (2021). *HEIMDALL. Multi-Hazard Cooperative Management Tool for Data Exchange, Response Planning and Scenario Building*. https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-17417/27631_read-71862

German Socio-economic Panel. (2023). *paneldata.org*. https://paneldata.org/

GESIS - Leibniz Institute for the Social Sciences. (2023). *GESIS Search*. https://search.gesis.org/

Google. (2023). *Google Crisis Response*. https://crisisresponse.google/

Hein, A., Schreieck, M., Riasanow, T., Setzke, D. S., Wiesche, M., Böhm, M., & Krcmar, H. (2020). Digital platform ecosystems. *Electronic Markets*, *30*(1), 87–98. https://doi.org/10.1007/s12525-019-00377-4

Hering, T., Karutz, H., Leuschner, V., Müller-Cyran, A., Neumann, O., & Zippert, T. (2021). *Psychosoziale Aspekte müssen im Corona-Krisenmanagement stärker berücksichtigt werden als bisher! [Psychosocial aspects must be given greater consideration in coronavirus crisis management than before!]*. https://www.asb-muenchen.de/application/files/5416/1546/2373/Positionspapier_PSNV_und_Corona-Krisenmanagement_2021-03-10.pdf.

International Organization for Migration. (2023). *Dashboard | Global Crisis Response Platform*. https://crisisresponse.iom.int/dashboard

Kilani, M. A., & Kobziev, V. (2016). An Overview of Research Methodology in Information System (IS). *OALib*, *03*(11), 1–9. https://doi.org/10.4236/oalib.1103126

Konnektable Technologies (2021). COncORDE v2.0 - Crisis Management Platform. https://konnektable.com/portfolio/concorde-crisis-management-platform/

Leidner, D. E., Pan, G., & Pan, S. L. (2009). The role of IT in crisis response: Lessons from the SARS and Asian Tsunami disasters. *The Journal of Strategic Information Systems*, *18*(2), 80–99. https://doi.org/10.1016/j.jsis.2009.05.001

Lüttschwager, N., Stelzmann, D., Gerhold, L., & Sterl, S. (2022). How Did Corona Crisis Managers in Germany Make Sense of the Psychosocial Situation? *European Journal for Security Research*, *7*(2), 163–189. https://doi.org/10.1007/s41125-022-00086-8

Martin-Lapoirie, D., McColl, K., Gallopel-Morvan, K., Arwidson, P., & Raude, J. (2024). Health protective behaviours during the COVID-19 pandemic: Risk adaptation or habituation? *Social Science & Medicine (1982)*, *342*, 116531. https://doi.org/10.1016/j.socscimed.2023.116531

Matheus, R., Janssen, M., & Maheshwari, D. (2020). Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities. *Government Information Quarterly*, *37*(3), 1–9. https://doi.org/10.1016/j.giq.2018.01.006

Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., Neuausgabe, 12., vollständig überarbeitete und aktualisierte Aufl.). *Beltz Pädagogik*. Beltz, J.

Mike Bostock and Observable. (2023). *D3: The JavaScript library for bespoke data visualization*. https://d3js.org/

Moberg, F., & Simonsen, S. H. (2011). *'What is resilience?'*. https://www.stockholmresilience.org/research/research-news/2015-02-19-what-is-resilience.html

Munich Security Conference. (2023). *Munich Security Index 2023: Munich Security Report.* https://securityconference.org/publikationen/munich-security-report-2023/munich-security-index-2023/

Netherlands Red Cross. (2023). *Community Risk Assessment: 510 Dashboards*. https://dashboard.510.global/#!/

Office for the Coordination of Humanitarian Affairs. (2023). *Natural Disasters Dashboard: Example: Afghanistan*. https://response.reliefweb.int/afghanistan/natural-disasters-dashboard

Otto, M., & Thornton, J. (2023). *Introduction bootstrap*. https://getbootstrap.com/docs/4.0/getting-started/introduction/

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Quinn, P., & Malgieri, G. (2021). The Difficulty of Defining Sensitive Data—The Concept of Sensitive Data in the EU Data Protection Framework. *German Law Journal*, *22*(8), 1583–1612. https://doi.org/10.1017/glj.2021.79

Red Cross Red Crescent. (2023). *IFRC GO*. https://go.ifrc.org/

Rogers, R. W., & Prentice-Dunn, S. (1997). Protection motivation theory. In D. S. Gochman (Ed.), *Handbook of Health Behavior Research I: Personal and social determinants.* 132). Plenum Press.

Schulze, A., Brand, F., Geppert, J., & Böl, G.-F. (2023). Digital dashboards visualizing public health data: A systematic review. *Frontiers in Public Health*, *11*, 999958. https://doi.org/10.3389/fpubh.2023.999958

Simon, H. (1969). *The Sciences of the Artificial.* MIT Press.

Smart Data. (2023). *Smart Data Disaster Management*. http://www.sd-kama.de/de/smart_data_disaster_management/#complete

Sterl, S., Billig, A., Taffo, F. W., & Gerhold, L. (2023). Visualizing the Psychosocial Situation in Crises and Disasters: Conceptualizing a Multi-Functional Crisis Information Platform (CIP-PS). In J. Radianti, I. Dokas, N. Lalone, & D. Khazanchi (Eds.), *Proceedings of the 20th International Conference on Information Systems for Crisis Response and Management, Proceedings of the 20th International Conference on Information Systems for Crisis Response and Management* (pp. 252–262). University of Nebraska at Omaha (USA). https://doi.org/10.59297/RRNM2516

Tandel, S. S., & Jamadar, A. (2018). Impact of Progressive Web Apps on Web App Development. *International*

Journal of Innovative Research in Science, Engineering and Technology, *7*(9), 9439–9444. https://www.researchgate.net/profile/Sayali-Tandel-2/publication/330834334_Impact_of_Progressive_Web_Apps_on_Web_App_Development/links/5c5605d3a6fdccd6b5dde018/Impact-of-Progressive-Web-Apps-on-Web-App-Development.pdf

Treffeisen, R., & Grosfeld, K. (2022). Herausforderung Klimawandel – Kommunikation und Wissenstransfer zwischen Fakten und gesellschaftlicher Handlungsnotwendigkeit. In M. Beuthner, U. Bomnüter, & J. A. Kantara (Eds.), *Risiken, Krisen, Konflikte* (pp. 237–255). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36195-2_14

United Nations Development Programme. (2020). *The Global Crisis Risk Dashboards*. https://data.undp.org/products/crisis-risk-dashboard

University of Erfurt, Bernhard Nocht Institute für Tropical Medicine, Robert Koch Institute, Federal Centre for Health Education, Leibniz Institute for Psychology, & Science Media Center. (2023). *Planetary Health Action Survey (PACE) - Explorer*. https://projekte.uni-erfurt.de/pace/explorer/

Weber, E. U. (2010). What shapes perceptions of climate change? *WIREs Climate Change*, *1*(3), 332–342. https://doi.org/10.1002/wcc.41

World Food Programme. (2023). *Emergency Dashboard: Example: Yemen*. https://www.wfp.org/publications/yemen

World Wide Web Consortium. (2012). *SKOS Simple Knowledge Organization System - home page*. https://www.w3.org/2004/02/skos/

Designing Information Flow with Diverse Community Health Workers to Improve Public Health Situational Awareness During COVID-19

Temitayo R. Okusanya
Medical College of Wisconsin
tokusanya@mcw.edu

Jane Gresser
Marquette University
jane.gresser@marquette.edu

Michael Stevenson
University of Wisconsin-Madison
mcstevenson@wisc.edu

Ajay Kumar
University of Münster
ajay.kumar@uni-munster.edu

MD Romael Haque
Marquette University
mdromael.haque@marquette.edu

Zeno Franco
Medical College of Wisconsin
zfranco@mcw.edu

ABSTRACT

Community health workers (CHWs) play a vital role in connecting communities, healthcare systems, and crisis response authorities. During a public health crisis, minority populations face increased disease and economic burdens, making adherence to health guidelines challenging. Despite numerous efforts to mitigate these impacts, the effectiveness is greatly compromised by communication gaps between communities and health authorities. Our study strategically employs the collaboration of CHWs and simple technology (Zoho Forms) to bridge gaps and enhance information flow between diverse hard-to-reach communities and public health authorities. We emphasize the unique contributions of CHWs to disaster response and the benefits of equipping CHWs with simple technology for effective communication during a public health crisis. Our work improved COVID-19 guidelines adherence, reducing infection spread while addressing critical mental health and resource needs within the communities served.

Keywords Community Health Workers (CHWs), COVID-19, Information Management, Public health Crisis Response, Simple Technology, Diverse Communities

INTRODUCTION

Efficient disaster management and response are dependent on effective information sharing and coordinated actions (Aung & Whittaker, 2012; Frassl et al., 2010; Iakovou & Douligeris, 2001)**.** These coordinated efforts require the quick assembling of response teams/systems made up of multiple cross-agency collaborations, to facilitate work distribution, increase response tempo, and, in public health crises, ensure the effective dissemination of critical information among diverse sections of the population (Castillo et al., 2018). Integrating community health workers (CHWs) as both key actors and information brokers between minority or marginalized communities and traditional public health response agencies is an important task during events like the recent COVID-19 pandemic. CHWs are uniquely positioned because they are embedded in communities that may be difficult for traditional response agencies to reach because of cultural, linguistic, and logistical barriers (Stevenson et al., 2022). However, CHWs are typically not considered part of the public health crisis response infrastructure, are not integrated into disaster planning, and are personally unfamiliar with communication and coordination protocols. A challenge faced by authorities during disasters is ensuring that the needs of individuals from minority populations/communities are adequately met as these populations are often reported to bear a greater burden during disasters (Hanson, 2021).

As of July 2023, Milwaukee County has reported 347,383 COVID-19 cases, and 2,532 deaths (USAFacts, 2023). At the onset of the pandemic, predominantly black and Latino communities in Wisconsin bore most of the burden of reported cases (Wisconsin Department of Health Services, 2020; The COVID Tracking Project, 2021). Studies conducted in Milwaukee County reported a high burden of COVID-19 infections among racially marginalized groups. In addition, several studies across the United States have consistently reported racial disparities in the incidence, mortality, and economic hardship related to COVID-19 (Alcendor, 2020; Park, 2021). In response to these disparities, the Milwaukee Health Department initiated a STOP-COVID-19 initiative in 2020 during the peak of the pandemic. The initiative was aimed at preventing the spread of COVID-19 in Wisconsin communities, particularly within racially diverse communities, and addressing the psychological needs of high-risk populations where traditional risk communication strategies may prove ineffective. As a result, a 3 -tiered strategy was employed to rapidly improve guidelines in these communities as well as meet their economic and psychological needs by 1) Increasing public awareness of key infection control information using culturally tailored messages, 2) Leveraging local media to implement culturally adapted micro-campaigns for risk reduction, and 3) Addressing psychological and wellness needs by leveraging community strength through CHWs. Identifying ways to optimize response operations with CHWs as a specific type of human-in-the-loop challenge offered specific practical, political, and ethical problems given the emphasis of the effort to ensure culturally adapted public health interventions in four distinct communities. Rapidly and equitably integrating CHWs into the Public Health response framework tested institutional abilities to manage high volume communication, making quickly deployed “off-the-shelf” or “lowest common denominator” technologies key to effective management of this process.

Here, findings are presented from a partnership involving the city of Milwaukee Health Department (MHD), the Medical College of Wisconsin (MCW), and Community Health Workers from four communities of color - African American, Hispanic, Hmong/SE Asian, and Native American working to rapidly form an ad-hoc organization to flexibly respond in culturally and linguistically adapted ways. These efforts required the rapid deployment of simple communication strategies to ensure that the frontline CHWs were able to move crucial, hyper-local information to traditional response agencies in a way that respected the unique contributions of these workers. Throughout this effort, rather than developing new technologies or deploying complicated mobile technologies, we focused on identifying off-the-shelf technology tools that were easily adapted and well understood by even the least technologically oriented personnel within the response team. For example, in other COVID-19 logistics efforts, we augmented Google Sheets by flowing information out of Sheets to SQLite for some processing, and then flowing the information back to the same Sheet so that all agency representatives could engage with the data (Franco et al, 2021). With the CHWs, we also focused on a lowest common denominator technology, in this case using Zoho Forms with some workflow logic to capture quick operational snapshots from each CHW that were then shared with the lead CHW for that community as well as with the Health Department and Medical School partners. The formal response agencies used this information to nuance the delivery of public health information, and food and financial relief; and focus staff and political actors' attention on the needs of these communities in ways that would have left the “on the ground " or residents’ voices out of the process.

PUBLIC HEALTH DISASTERS

The onset of an infectious disease outbreak is considered a public health emergency, necessitating the establishment of new protocols and the implementation of strategies and methodologies to address the swift changes caused by the outbreak (Maciel et al., 2020). The consequences of such outbreaks are highly potent and extend beyond the infected individual(s), subsequently encompassing the entire population, economies, public health systems, and social interaction (Hu & Qiu, 2020). An example is seen in the most recent infection outbreak (COVID-19), originating in Wuhan City, China in December 2019, resulting in widespread fatalities, reduced quality of life, economic hardship, reduced global trade and travel, as well as constrained social interactions (Bhaumik et al., 2020). Within the context of disease outbreaks of high magnitude like the COVID-19, Ebola, and H1N1 viruses, effective emergency management hinging on efficient systems and the harmonious working together of both internal and collaborative governmental bodies are essential components for implementing containment measures within the framework of emergency management (Hu & Qiu, 2020). Some of the key contributors to this collaborative emergency management approach include the state Department of Health, hospitals, community clinics, NGOs, businesses, healthcare workers (including doctors, nurses, microbiologists, and CHWs), and most importantly the individuals that make up the population (Palafox et al., 2020).

Disaster Communications

During disasters, effective risk communication is vital to addressing public fears, stigmas, informed decision-making, and adherence to established guidelines (Porat et al., 2020). Risk communication and community engagement (RCCE) are those processes and methodologies used to systematically consult and dialogue with communities at most risk of infection, or those whose cultural, religious, or ethical practices impact risk (Adebisi et al., 2021). Furthermore, the chains of communication during disasters are multifaceted, for example, during the COVID-19 pandemic, the World Health Organization's (W.H.O.) guidelines on risk communication encouraged a two-way dialogue/communication pattern with communities, the public, and stakeholders that ensure that information is accurately tailored to their specific circumstances (World Health Organization, 2020). In addition to this, a community-based participatory approach (CBPR) is required, one that allows individuals free access to information on potential risks and provides them with the right to participate in the disaster response process (Nylund, 2022).

COMMUNITY HEALTH WORKERS (CHWs)

According to the International Labor Organization (ILO), community health workers (CHWs) play a vital role in providing health education, support, and referrals encompassing diverse services. Their responsibilities extend to assisting individuals, families, and communities in adopting preventive health measures and facilitating access to appropriate healthcare and social services (Bezbaruah et al., 2021). Particularly, CHWs drawn from marginalized communities occupy a unique position as trusted sources of health information (Ayodele et al., 2024). Their status as members of the communities they serve grants them invaluable insight into cultural norms and health-belief models. The credibility of CHWs serves as a crucial link between communities and healthcare organizations, wrestling in heightened community health literacy and improved access to essential resources through their coordination efforts (Wennerstrom et al., 2013; Scheib & M. Brinton Lykes, 2013).

CHWs have received increasing attention as they may play a pivotal and distinct role in the disaster response framework (Nicholls et al., 2014). The Centers for Disease Control and Prevention (CDC) defines CHWs as "frontline workers who are trusted members or individuals who have a good understanding of the communities they serve" (CDC, 2022). CHWs are an essential part of public health disaster management processes, serving as a bridge between communities and the resources they need. Their main responsibilities involve building trusting relationships between communities and stakeholders, eliminating barriers to access, and lending a voice to issues affecting the communities they serve (U.S. Department of Health & Human Services, 2019). The inclusion of CHWs in disaster management, however beneficial, is plagued with its unique challenges in that CHWs have limited knowledge of official language, might have trust issues with large institutions same as the population they serve, and are oftentimes not considered as formal employees. In addition, CHWs differ from traditional medical or public health crisis responders as they bring a different skill set, community social capital/relationships/connections, and responsibilities to the crisis response team (Pinto et al., 2020).

Role of CHWs in Public Health Disaster Management

During disease outbreaks of high magnitude such as the COVID-19, Ebola, and H1N1 viruses, the alignment of both internal and collaborative governmental bodies are essential components for implementing containment measures within the framework of emergency management (Hu & Qiu, 2020). While numerous studies have demonstrated the efficacy of CHWs in reducing the burden of chronic diseases such as diabetes and hypertension in minority and underserved populations (Nassar et al., 2013; O'Connor et al., 2020), a lack of information exists on the use, management, and information flow between CHWs and public health systems in widespread disaster events, such as a pandemic. Even less is known about how CHWs interact with technology in the context of crisis events.

The importance of CHWs in disaster preparedness stems from the personal ties they have to the communities they serve, and thus their ability to educate and recruit help, monitor the effectiveness of interventions, and fill in gaps in health service (Boyce & Katz, 2019). Previously, CHWs were mobilized to identify the healthcare and social needs of individuals during Hurricane Sandy. CHWs were effective in setting health goals for patients with chronic medical conditions, identifying community resources, and counseling patients (Russell et al., 2018). Similarly, during the 2017 Peru floods, CHWs were trained to support mental health interventions by conducting screenings for depression and domestic violence and connecting individuals to the necessary support services (Contreras et al., 2018). They also implemented activities that allowed children to relieve stress while building emotional management skills and forums to encourage community participation. The post-disaster interventions exposed the inadequacies of the health systems in addressing the needs of those marginalized communities and the role CHWs can play in bridging those gaps.

Employing CHWs provides several key operational advantages within a public health disaster management framework. Based on our observations working with about 12 CHWs as a fundamental institutional healthcare approach to managing COVID-19 (Franco et al., 2023; Stevenson, 2023), we found several essential characteristics exhibited by CHWs when viewed through the lens of crisis logistics, actions, information flow, and situational awareness:

1. **Autonomous by nature** - Each CHW functions as a largely autonomous agent within their community, operating independently throughout most of their work, only occasionally checking in to provide information and take directives. While this is both good and bad - if carefully selected, they don't require much instruction about what to do and can make independent decisions without creating a burden for these decisions on traditional crisis response managers for most tasks. On the other hand, because the crisis is rapidly unfolding, and their check-ins are relatively infrequent there may be substantial delays in information transfer, redirection of high-level activities, etc (Malcarney et al., 2017).
2. **Hub and Spoke** - Outside of a crisis context, CHWs typically operate in a hub and spoke model of organization, with each CHW reporting to a central command element independently – this model is a key strength, as it relies on high levels of trust between CHWs and the command element from the healthcare system. This relationship is, in a perfect world, fairly mutual – such that the CHW is regarded with respect even though they are very different from the traditional healthcare system. This respect is demonstrated through non-hierarchical access directly into the healthcare system.
3. **High Community Trust** - Benefiting from high levels of community trust, CHWs possess valuable insight into the community's experiences and needs. This trusted position enables them to identify areas where the public health system may fall short in addressing the community's needs.
4. **High Quality, Difficult to Obtain Information Flow** - Because the community trusts them, they can provide direct, unfettered views of things on the ground (remote sensing) that can be sent back to the public health agency in the form of text message updates, photographs, brief videos (Baker et al., 2024, Taylor et al., 2017).
5. **Culturally Competent Information Flow & Sense Making** - CHWs are excellent at translating or transliterating complex, linguistically, and culturally opaque information for the public health response agency. They have a unique ability to convey intricate details in a manner that aligns with both public health agencies and the community's language and cultural context, thereby enhancing the effectiveness of information flow and sense-making in response efforts (Griffith et al., 2023).
6. **Trusted Asset Logistician** - CHWs can distribute essential supplies like personal protective equipment (PPE), food, diapers, baby formula, and monetary donations or gift cards to individual households that had COVID-19 positive household members and needed to isolate to prevent workplace infection

spread and were in financial need - in part because they have were trusted with these tasks before the pandemic, and because they have hyper-local networks through which the assets can be channeled, etc.

However, despite their numerous benefits, CHWs frequently encounter challenges while performing their duties, some of which include:

1. **Excessive Autonomy:** Sometimes, the heightened level of autonomy CHWs possess raises concern that their initiatives could run counter to or do not assist the public health agencies' primary goals.
2. **Loss of Shared Situational Awareness in transition to Hierarchical Structure** – Even though we wanted to preserve the hub and spoke model and they are largely autonomous, the complexity of the COVID-19 response meant that the formal response structures were rapidly overwhelmed by information and asset requests from the CHWs. We had to switch to a hierarchical model in a respectful, culturally competent way. However, this meant that we then lost direct insights into what was happening on the ground because the information from most of the CHWs was being filtered through a lead CHW for each community — this became in the crisis response literature a "shared situational awareness" problem. Critical information available to a CHW lower in the hierarchy might be filtered out before reaching the Public Health Department.
3. **Coordination Challenges**- Coordination problems arise in the identification and resolution of needs across diverse communities. This often involved understanding the intersections of hyperlocal networks among CHWs and determining how these shared networks can be leveraged for more effective collaboration and response, while also respecting boundaries and not moving past the level of trust the CHW developed.
4. **Safety and security of CHWs in the field:** In utilizing the help of CHWs for crisis response, there is a paramount concern around guaranteeing their safety and security in the field.
5. **Mental Health of CHWs:** Additional concerns also revolve around acknowledging and addressing the mental health and well-being of CHWs as they perform these crucial tasks (Johnson et al., 2022).

Balancing the multifaceted role of CHWs in public health disaster management necessitates striking a balance between the strengths and weakness of the CHW model at the interface with formal response structures, thereby necessitating a need for cultural tailoring, strengthened supervisory systems, and a continuous communication and feedback loop to ensure that the positive impacts of CHWs in disaster management prevail while addressing potential pitfalls to maximize their effectiveness.

OFF-THE-SHELF HEALTH INFORMATION SYSTEMS IN DISASTER MANAGEMENT

Over the years, the integration of technological tools in disaster management has yielded significant benefits. These tools have been used in forecasting, hazard monitoring, disaster mapping, emergency response, and information coordination, all of which collectively contribute to a comprehensive framework for addressing and mitigating the impact and damages caused by disasters (Rathore, 2016; Eguchi et al., 2008). Dating back decades, major public health disasters have utilized the strategic deployment of advanced technological tools as an intricate component for effective response and recovery efforts exemplified by historical events such as Hurricane Andrew in 1992, and the 1994 Northridge earthquake. During these disasters, Geographical Information Systems (GIS) played a pivotal role by providing visual and geospatial information for efficient response. Similarly, the aftermath of the World Trade Center attacks saw the utilization of technological tools for conducting damage assessment, and facilitating recovery operations (Eguchi et al., 2008). In more recent instances, the Ebola outbreak in Sub-Saharan Africa, and the ongoing COVID-19 pandemic have further underscored the benefits of technological tools in designing, planning, and implementing best practices for infection control, the provision of response to affected individuals and communities (Whitelaw et al., 2020).

While many of these events have reported the use of sophisticated and advanced technology, simple off-the-shelf technology has also been beneficial for disaster management, most importantly for disaster communication and management (Gómez, 2010). A study led by Kavota et al., 2020 analyzing social media adoption and its usefulness in disaster management found that user adoption was driven by perceived ease of use and advantages including data accessibility and flexibility. Further, their study concluded that of the approximately 3.848 billion active social media users in the world, 3.256 billion of these access social media via their smartphones (Kavota et al.,

2020). Other key advantages of using social media during disasters include bi-directional communication, increased awareness of rapidly changing conditions, and geolocation of users. For example, people impacted by Hurricane Sandy in 2012 posted upwards of 800,000 hurricane images using #Sandy, each of which documented visual evidence of changing conditions and showed the geographic location where the picture was taken (Sakurai & Murayama, 2019). These events both summarize the impactful role of simple technology and underscore the significance of accessible tools, and social networks in disaster communication and management.

METHODOLOGY

This research is presented as a practice based, in-depth case study (Kuhlicke, 2013; Kankanamge, Yigitcanlar & Goonetilleke, 2020). When the COVID-19 crisis struck, we rapidly shifted toward a response footing, interestingly moving some of us with extensive crisis research backgrounds, but little practical crisis response experience in responder roles. Thus, the case study is grounded in an action research approach (Landgren, 2010; Lempinen & Rajala, 2014). Because the authors were forced to switch roles from researcher to responder, we were in effect designing information systems on the fly to solve problems as they came up (Cross, 1982). Wearing both hats, the response work took precedence. Data collection was therefore primarily opportunistic instead of dictated by a predetermined research question or hypothesis. By shifting back to a research perspective and reconstructing the complete case study after the fact to be reported to the crisis management research community, we were allowed an opportunity to reflect and capture as much of this process and the learnings from it as possible. This event reconstruction approach allowed us to pause and understand these additional elements for the first time as a group, and we worked to ensure that each of the authors on the response team was satisfied that the telling of the events accurately represented their understanding within the overarching effort as best as possible. This approach afforded us a much more comprehensive view of the actions taken and paints a picture of the complexity encountered by response teams in completely improvised settings (Zhang & Mendonça, 2020). Notably, this perspective includes only the healthcare and public health agency perspectives, not those of the CHWs, which are presented elsewhere (Stevenson, et al. 2023).

RESULTS

The onset of the Covid-19 pandemic brought multiple challenges to CHWs including the need for quick action, ever-evolving information, and large-scale impact throughout the community. Our previous work with CHWs used a dedicated supervision model, however, the COVID-19 pandemic required a significantly larger reach and therefore we moved to a cascade system.

STOP COVID-19 Community Health Workers

During COVID-19, we sought to provide CHW support to four communities of color in Southeastern Wisconsin, specifically within African American, Hispanic, Hmong/SE Asian, and Native American communities. The overall STOP COVID-19 project and the CHW program embedded within this public health crisis response effort are detailed elsewhere (Franco et al., 2023; Stevenson et al., 2023). Each of the four communities of color served was designated one CHW lead and 1-3 additional CHWs reporting to each lead. As such, technology was recognized as a critical means of communication between our 4 teams of CHWs, their respective team leads, and project leadership. In this project, CHWs were short-term employees of the City of Milwaukee through the STOP COVID-19 grant, and data provided here was administratively gathered as part of this paid role. The information was used to manage the overall crisis response.

Zoho CHW Impact Form

The form used in the STOP COVID-19 Initiative was labeled “CHW Impact Form”, this allowed all CHWs and Team Leads to submit weekly progress reports and issues/ needs (DS Table 1). Zoho online forms utilize a ‘drag and drop’ builder allowing for rapid creation and customization of the form. The data entry burden was minimized using checkboxes and predefined options that were defined through rapid collaborative design processes (Klein et al., 2003) with the CHWs, while also providing several areas for free-form entry. At the onset of the project the form captured immediate issues, agencies served, family services provided, community education/ outreach,

and team activities. Following vaccine availability in February 2021, additional questions were added to capture vaccine compliance, concerns, and education within the four communities.

Although not unique to Zoho Forms, we selected this option over more basic tools like Google Forms because Zoho has built-in form logic, which allows data elements used in the form to act on form submission. Upon form submission, a PDF report was automatically emailed to the respective CHW Team lead and project leadership at MHD and MCW. This ensured direct communication by CHWs reached all levels of leadership unfiltered, and facilitated response to community needs while reducing overall volume of communication. While the same types of conditional logic can be accomplished through other projects or scripting language augments to Google Forms, we had limited access to a programmer and elected to use the programmer's time for a more complicated logistics management effort (Franco, et al., 2021).

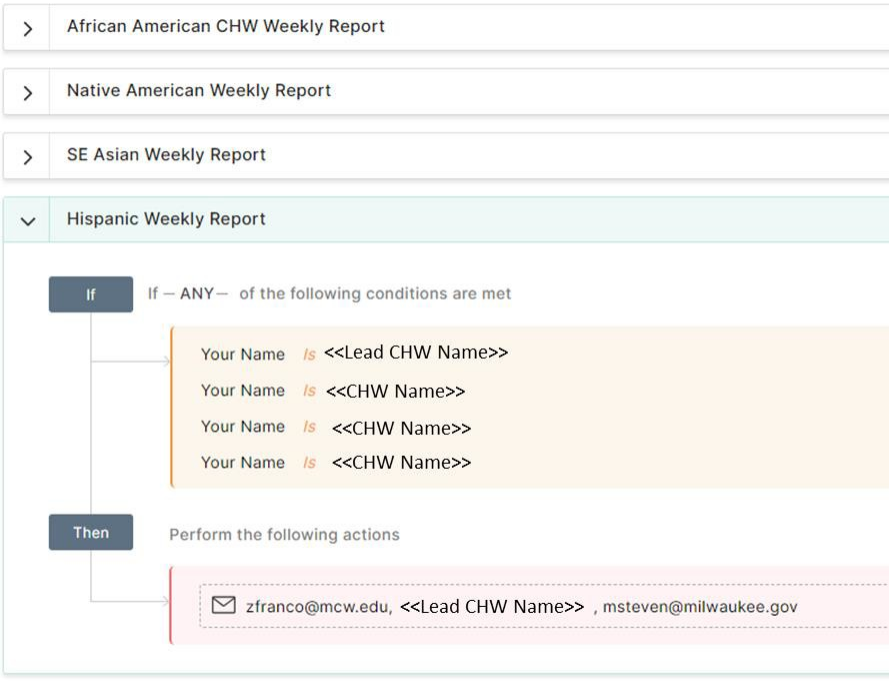


Figure 1. Data Entry by CHWs Automatically Triggers &Generates Situational Reports Sent to Leadership at MHD and MCW via Zoho Forms.

CHW Social Information Context & Overwhelm

Notably, we observed that the CHWs were operating in a very different information environment, working directly with residents in their neighborhoods through things like private Facebook groups, individual and group SMS Text, and Facebook Live to broadcast information about things like food pantries (notably these FB live events were often narrated in the CHWs primary language), and older technology approaches like phone trees and phone check-in lists. One of the CHWs reported that their team was contacting hundreds of households a week by voice phone call check-in. Importantly, the Zoho form approach created a single channel of information for the command layer of this project from the CHWs. The intensity of communication from the lead CHWs, who had direct phone and text access to the project managers, alone threatened to overwhelm the command layer (Rao et al. 2017, Misra et al. 2020). We were forced to push most synchronous upward communication from the CHWs to the lead CHW for their population group, adhering to a hierarchical model. However, the Zoho form allowed unfiltered information from the CHWs to still reach the program staff and be evaluated and actioned asynchronously. In addition to the Zoho forms, a Facebook group for all CHWs (regardless of population served), lead CHWs and program staff to view operationally related photos asynchronously.

Responses

CHWs submitted 150 responses to the form over a one-year period between August 2020 and August 2021 spanning the height of the pandemic into the early vaccination and recovery phase. Hispanic CHWs returned 62 community updates, African American CHWs 34, Native American 42, and SE Asian 12. The low numbers for the SE Asian community likely reflect the tenuous connection we had to this community beyond the lead CHW, although information provided by three SE Asian CHWs was provided.

COMMUNICATION FLOW BETWEEN CHWs AND STAKEHOLDERS:

CHWs were asked to complete the Impact Form at least once a week, whereupon submission, the form was sent automatically to CHW Leaders, the Public Health Department, and MCW (see Figure 1).

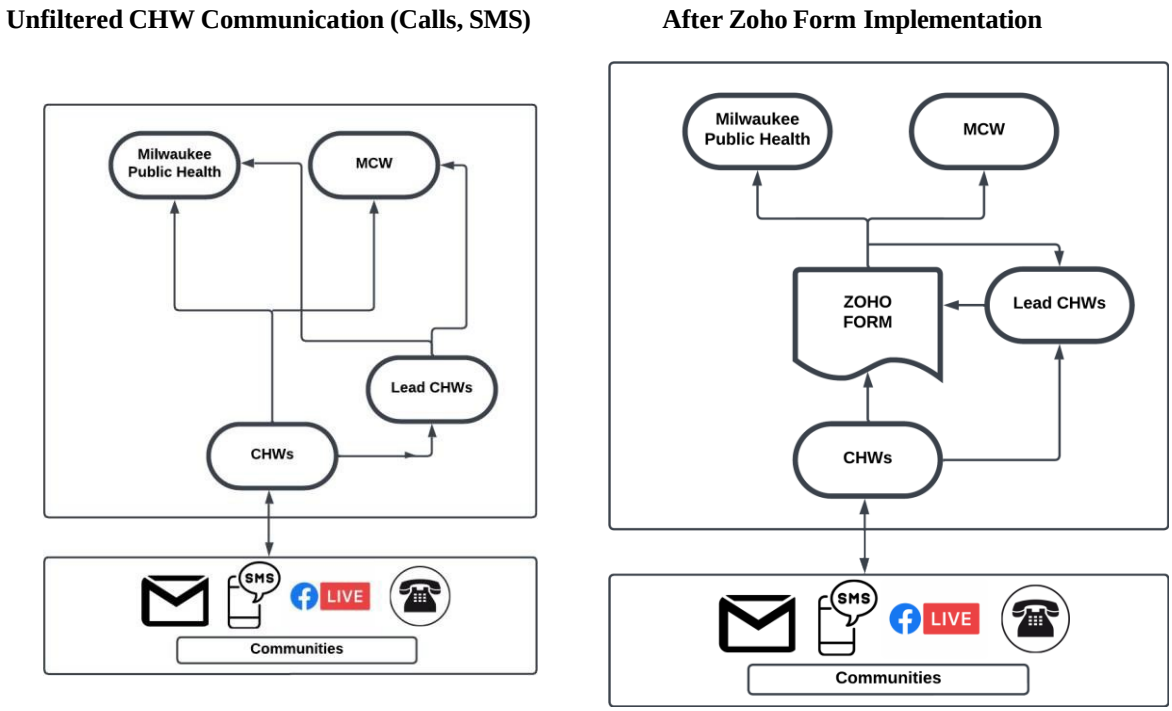


Figure 2. Weekly updates using simple ZOHO Form and submitted automatically to Milwaukee Public Health, Medical College of Wisconsin (MCW), and the Lead Community HealthCare Workers.

Below, we provide a detailed summary of key efforts implemented across these communities based on information obtained from Zoho Forms.

DS Table 1. CHW-Generated Reports from Zoho Forms for Agencies Served

ZOHO FORM DATA FIELD	EXEMPLAR INFORMATION
Agencies Served	
Agency Types Served this Week? (check all that apply): Church, Social Service agency, Clinic, Food Pantry, Homeless Shelter, School, Other - Please describe	**Other agencies reported included:** • Community leadership programs • Other Native American Tribes (e.g. • Stockbridge-Munsee Tribe) • Home Health agencies • Local newspapers • Financial services companies • Check cashing locations • Retail cell phone stores • Culturally specific food stores • Cultural/spiritual events (Unity Fire, Sobriety Fire) • Consulates • Voting drives • Toy drives
Notes on Agencies served/ Number served (free text) *Any impact information you think would be helpful for us to know about, questions that came up from agencies, etc?*	**Native American (09/04/2020)** *I dropped off a total of over 300 masks at the Potawatomi Foundation Leadership Program and the Southeastern Oneida Tribal Services, as well as some posters and COVID-19 information. At the Water Walk Milwaukee 2020 I passed out masks and cards with COVID-19 information. Finally, at the Natives in MKE Photo Campaign 2020, I passed out masks, distributed food, and gave out support resources.* **African American (09/07/2020)** *Met with NAACP and discussed the distribution of masks and resource information at a voter registration event 9/19/20.* *Met with the American Legion Post in Kenosha to assist in distributing masks at an event in Kenosha.* *delivered Pampers to Mexican Fiesta, along with masks to <Native American CHW>* *Contacted United Way for essential COVID-19-related supplies, United Way will be delivering items to Dryhootch on Brady St. hygiene kits, (diapers, masks, sanitizer, and other essentials) I will be contacting CHW leads for distribution of items this week. Working with Feeding America to process food orders for CWH leads.*

DS Table 2. CHW-Generated Reports from Zoho Forms for Families Served

ZOHO FORM DATA FIELD	EXEMPLAR INFORMATION
Families Served	
Types of Family Services Provided This Week? (check all that apply): COVID-19 information, Mental Health Resources, Food, Masks, Household Goods, Advice/ Personal Support, Referral to government services (e.g. rent, support etc.), Support to Families with COVID-19 positive in households, Phone calls with families, Coordinating with other CHWs or agencies, Other - Please specify	• Transportation assistance • Christmas gifts for children • CHW generated live stream family support event. • Assistance editing job resumes. • Diapers, wipes formula • Referral to immigration, criminal, and family lawyers • Cancer screening • Feminine sanitary products
Notes on families served/ number served (free text) Any impact information you think would be helpful for us to know about, questions that came up from families, etc?	Hmong Community *The Hmong community is having to cope and grieve non-traditionally and I am coping a family through losing a loved one. Also, I am working with families to get them access to food pantries and culturally specific foods.* *09/11/2020* *I am continuing to work with a teen who lost his dad due to COVID-19. The entire family is having a hard time coping and I am making referrals and being another source of support. I am also receiving more food donations. I am individually dropping them off at family's homes.* *09/18/2020* *I am doing regular check-ins with a family. The mom is on her way to recovery but they mentioned to me how even after COVID, things do not go back to normal. The community puts a stigma on them as if catching coronavirus was a curse. People are afraid of their families even after they have recovered from COVID-19. She has physical limitations that were not there before covid such as tiredness from just going up the stairs, breathing still hurts, and physical tiredness. The youth I am working with is finding support through opening up to friends and family again. I think sometimes talking to close friends and family isn't always the best option, so talking to me has been helpful.*

DS Table 3. CHW-Generated Reports from Zoho Forms on Community Empowerment

ZOHO FORM DATA FIELD	EXEMPLAR INFORMATION
Community Education/ Outreach/ Other	
Notes on Community Education, Outreach, and Coordination (free text)	
Team and Learning Activities	
Types of team activities? (check all that apply): Assisted another team member, attend a conference/ presentation, Attended a training	
Notes on Team and Learning Activities (free text)	**Hmong (9/18/2020)** <Native American CHW lead> needed another mask for this cultural art project and so I assisted him in that. I made masks at Mexican Fiesta with <Hispanic CHW Lead), <Native American CHW>, and <Hispanic CHW>. I also met with <Hispanic and African American CHW leads> at <Organization Name> and picked up some supplies for distribution. <African American CHW lead> has been helpful in getting materials to me and the team. My team and I are getting ready to put kits together to distribute. Kits will include Health Department masks, hand sanitizer, latex gloves, and the COVID info card.

Mental Health Support: A MH webinar was conducted across all 4 communities to provide training to CHWs enabling them to act as a MH first response worker within their communities. Guidance on accessing MH support, and referrals to government-owned facilities were also provided to the CHWs, hence equipping them with the necessary tools to address the MH crisis on the go.

COVID-19 Support: CHWs were able to identify resource gaps within their communities that could prevent infection spread, and through their active reporting via the Zoho forms, efforts were made to distribute face masks, provide tailored information on COVID-19 testing and masking protocols across all 4 communities. The CHWs also created awareness and dispelled rumors and negative societal stigma associated with the virus.

Facilitation of Local Community Support Teams/ Systems: Aside from their commitment to reporting first-hand information to leadership (MCW, Milwaukee Public Health Department & CHW Leads), the CHWs through the data collected helped in facilitating outreach work between local churches and the communities they serve. As a result of these efforts, churches, and other non-profits were able to meet the resource needs of individuals financially constrained due to the pandemic by providing food pantries, rental support, clothing, and hygiene supplies.

DISCUSSION

Our study provides an in-depth narrative of the pivotal role CHWs play in navigating the complex terrain of public health crisis response, particularly within diverse communities. We have underscored the role of communication and information management during public health crises while highlighting the various tiers of partnerships and

communication efforts employed by health departments to obtain information from hard-to-reach populations. Adding to the significance of CHWs in public health crisis response, a distinction to our approach is the strategic integration of simple technology as illustrated by the use of Zoho forms in our study. Our approach emphasizes the need to look beyond conventional technological solutions requiring expertise and infrastructure for information management during a crisis. The dynamic integration of CHWs and Zoho forms during the Stop-COVID-19 efforts across four diverse communities in Milwaukee proved beneficial for on-the-go crisis management efforts. Equipping the CHWs with this simple technology was a crisis response strategy that transcended the barriers of complex tech adoption. Furthermore, in keeping the tools simple, we were able to empower CHWs and streamline information flow thereby establishing a robust infrastructure for the disaster response frameworks within these diverse communities.

While the simplicity of the technology used in our study was a strength of our COVID-19 crisis response efforts, we however acknowledge the potential challenges that may arise concerning data security and privacy (Kim et al., 2017). In terms of data privacy and security concerns, while Zoho forms as a security component to ensure data safety and privacy, we however, think it is additional efforts must be made to ensure CHWs are informed on additional measures they must take to ensure the data they collect within the communities they serve are properly secured, in addition to this, CHWs must also assure communities from which this information are obtained about the safety of their health information to avoid mistrust and potential loss of the life-long partnership already established between CHWs and the communities they serve.

CONCLUSION

In healthcare systems, Community health workers have been proven to provide substantial and indispensable support for routine care activities, disease management, and infection prevention during public health crises. Including their roles with additional support for crisis reporting and information management will further enhance crisis management efforts by public health authorities. Our study, centered on leveraging the use of Zoho forms as a simple technology during the Stop COVID-19 efforts in collaboration with Milwaukee's public health department, academic partners at the Medical College of Wisconsin, and community leaders in tandem with CHWs, demonstrated significant benefits in communities mostly impacted and in dire need of resources. While we can say that our study successfully met its objectives within the target communities, we, however, emphasize the need for a continuous commitment to identifying and enhancing public health communication and information management within these communities during public health crises and beyond. We emphasize the importance of leveraging CHWs and empowering them with simple technologies like Zoho forms, to facilitate their work. Further work should begin to explore strategies to consolidate the role of CHWs as essential contributors to public health crisis response frameworks. In addition, incorporating user-friendly and quick reporting tools for information management, reporting, and distribution, while upholding privacy, ease of use, and confidentiality should be considerations for their uptake and impact.

Our study advocates for a continuous commitment to securing and strengthening the role of CHWs and integrating accessible technologies into the frameworks of efficient public health crisis management.

ACKNOWLEDGEMENT

This work was wholly supported by the Advancing a Healthier Wisconsin COVID Rapid Response Grant, *Stop COVID-19.*

REFERENCES

Adebisi, Y. A., Rabe, A., & Lucero-Prisno III, D. E. (2021). Risk communication and community engagement strategies for COVID-19 in 13 African countries. *Health Promotion Perspectives*, *11*(2), 137–147. https://doi.org/10.34172/hpp.2021.18

Akintobi, T. H. (2020). Community Engagement of African Americans in the Era of COVID-19: Considerations, Challenges, Implications, and Recommendations for Public Health. *Preventing Chronic Disease*, *17*. https://doi.org/10.5888/pcd17.200255

Alcendor, D. J. (2020). Racial Disparities-Associated COVID-19 Mortality among Minority Populations in the US. *Journal of Clinical Medicine*, *9*(8), 2442. https://doi.org/10.3390/jcm9082442

Aung, E., & Whittaker, M. (2012). Preparing routine health information systems for immediate health responses to disasters. *Health Policy and Planning*, *28*(5), 495–507. https://doi.org/10.1093/heapol/czs081

Ayodele, James O., Marika L. Kromberg Underwood, Duaa Al Ammari, Kara Goldstone, and Emmanuel Agogo. "Enhancing trust and transparency for public health programs." In Modernizing Global Health Security to Prevent, Detect, and Respond, pp. 457-473. Academic Press, 2024.

Baker, N. D., Franco, Z., & Okusanya, T. (2024, February 18). Sensorized health interventions in marginalized communities: The role of surveillant assemblages in maintaining compliance within disaster-affected healthcare systems. The 8th IFIP WG5.15 Conference on Information Technology in Disaster Risk Reduction (ITDRR2023). https://www.academia.edu/115082427/Sensorized_health_interventions_in_marginalized_communities_the_role_of_surveillant_assemblages_in_maintaining_compliance_within_disaster_affected_healthcare_systems

Bezbaruah, S., Wallace, P., Zakoji, M., Padmini Perera, W. S., & Kato, M. (2021). Roles of community health workers in advancing health security and resilient health systems: emerging lessons from the COVID-19 response in the South-East Asia Region. *WHO South-East Asia Journal of Public Health*, *10*(3), 41. https://doi.org/10.4103/2224-3151.309872

Bhaumik, S., Moola, S., Tyagi, J., Nambiar, D., & Kakoti, M. (2020). Community health workers for pandemic response: a rapid evidence synthesis. *BMJ Global Health*, *5*(6), e002769. https://doi.org/10.1136/bmjgh-2020-002769

Boyce, M. R., & Katz, R. (2019). Community Health Workers and Pandemic Preparedness: Current and Prospective Roles. *Frontiers in Public Health*, *7*. https://doi.org/10.3389/fpubh.2019.00062

Castillo, A., Reyes, N., & Luz Esperanza Bohórquez. (2018). Organizational Features in Disaster Risk Management Systems. 155–174. https://doi.org/10.1007/978-3-030-32169-7_12

CDC. (2022, February 18). Community Health Worker Resources | CDC. Www.cdc.gov. https://www.cdc.gov/chronicdisease/center/community-health-worker-resources.html

Contreras, C., Aguilar, M., Eappen, B., Guzmán, C., Carrasco, P., Millones, A. K., & Galea, J. T. (2018). Community strengthening and mental health system linking after flooding in two informal human settlements in Peru: a model for small-scale disaster response. Global Mental Health, 5. https://doi.org/10.1017/gmh.2017.33

Eguchi, R., Huyck, C., Ghosh, S., & Adams, B. (20). The Application of Remote Sensing Technologies for Disaster Management.

Franco, Z., Davis, C., Kalet, A., Horng, M., Horng, J., Hernandez, C., Dotson, K., Yaspan, A., Kumar, A. and Lijnse, B., 2021. Augmenting Google Sheets to Improvise Community COVID-19 Mask Distribution. In 18th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2021 (pp. 359-375).

Franco, Z., Davis, C.S., Kalet, A., Hooyer, K., Nelson, D., Amin, Q.E., Stevenson, M., Cox, K., Yaspan, A., Perkins, H. and Kryshak, T., 2023. Medical School Civic Engagement During COVID-19: Activating Institutions for Equitable Community Response. Journal of Humanistic Psychology, p.00221678231206202.

Frassl, M., Lichtenstern, M., Khider, M., & Angermann, M. (2010). Frassl et al. Developing a System for Information Management in Disaster Relief Developing a System for Information Management in Disaster Relief -Methodology and Requirements. https://elib.dlr.de/63783/1/229-Frassl_etal.pdf

Gómez, E. (2010). The Role of SMS Text Messaging to Improve Public Health Response. Chapman and Hall/CRC EBooks. https://doi.org/10.1201/b10315-8

Griffith, D. M., Efird, C. R., Baskin, M. L., Webb Hooper, M., Davis, R. E., & Resnicow, K. (2023). Cultural sensitivity and cultural tailoring: Lessons learned and refinements after two decades of incorporating culture in health communication research. Annual Review of Public Health, 45(1). https://doi.org/10.1146/annurev-publhealth-060722-031158

Hanson, M. (2021, May 21). Inoculation Nation: Limited COVID-19 vaccine data shows uneven access by race. APM Research Lab. https://www.apmresearchlab.org/covid/vaccines-by-race

Hu, G., & Qiu, W. (2020). From guidance to practice: Promoting risk communication and community engagement for prevention and control of coronavirus disease (COVID-19) outbreak in China. Journal of Evidence-Based Medicine. https://doi.org/10.1111/jebm.12387

Iakovou, E., & Douligeris, C. (2001). An information management system for the emergency management of hurricane disasters. International Journal of Risk Assessment and Management, 2(3/4), 243. https://doi.org/10.1504/ijram.2001.001508

Johnson, L. J., Schopp, L. H., Waggie, F., & Frantz, J. M. (2022). Challenges experienced by community health workers and their motivation to attend a self-management programme. African Journal of Primary Health Care & Family Medicine, 14(1). https://doi.org/10.4102/phcfm.v14i1.2911

Kankanamge, N., Yigitcanlar, T., & Goonetilleke, A. (2020). How engaging are disaster management related social media channels? The case of Australian state emergency organisations. International Journal of Disaster Risk Reduction, 48, 101571.

Kavota, J. K., Kamdjoug, J. R., & Wamba, S. F. (2020). Social Media and disaster management: Case of the North and South Kivu regions in the Democratic Republic of the Congo. International Journal of Information Management, 52, 102068. https://doi.org/10.1016/j.ijinfomgt.2020.102068

Kim, B., Johnson, K., & Park, S.-Y. (2017). Lessons from the five data breaches: Analyzing framed crisis response strategies and crisis severity. Cogent Business & Management, 4(1), 1354525. https://doi.org/10.1080/23311975.2017.1354525

Klein, M., Sayama, H., Faratin, P., & Bar-Yam, Y. (2003). The dynamics of collaborative design: insights from complex systems and negotiation research. *Concurrent Engineering*, *11*(3), 201-209. https://doi.org/10.1177/10632930303802

Kuhlicke, C. (2013). Resilience: a capacity and a myth: findings from an in-depth case study in disaster management research. Natural hazards, 67(1), 61-76

Maciel, F. B. M., Santos, H. L. P. C. dos, Carneiro, R. A. da S., Souza, E. A. de, Prado, N. M. de B. L., Teixeira, C. F. de S., Maciel, F. B. M., Santos, H. L. P. C. dos, Carneiro, R. A. da S., Souza, E. A. de, Prado, N. M. de B. L., & Teixeira, C. F. de S. (2020). Community health workers: reflections on the health work process in Covid-19 pandemic times. Ciência & Saúde Coletiva, 25, 4185–4195. https://doi.org/10.1590/1413-812320202510.2.28102020

Malcarney, M., Pittman, P., Quigley, L., Horton, K., & Seiler, N. (2017). The changing roles of Community Health Workers. Health Services Research, 52(S1), 360–382. https://doi.org/10.1111/1475-6773.12657

Misra, S., Roberts, P., & Rhodes, M. (2020). Information overload, stress, and emergency managerial thinking. International Journal of Disaster Risk Reduction, 51, 101762. https://doi.org/10.1016/j.ijdrr.2020.101762

Monte, L., & Perez-Lopez, D. (2021, July 21). How the Pandemic Affected Black and White Households. The United States Census Bureau. https://www.census.gov/library/stories/2021/07/how-pandemic-affected-black-and-white-households.html

Nassar, A. F., Alemi, F., Hetmyer, A., Alemi, Y., Randolph, L. A., & Ramey, S. L. (2013). Automated Monitoring to Detect H1N1 Symptoms Among Urban, Medicaid-Eligible, Pregnant Women: A Community-Partnered Randomized Controlled Trial. Journal of Community Health, 39(1), 159–166. https://doi.org/10.1007/s10900-013-9754-1

Nicholls, K., Picou, J. S., & Lowman, J. (2014). Enhancing the Utility of Community Health Workers in Disaster Preparedness, Resiliency, and Recovery. International Oil Spill Conference Proceedings, 2014(1), 170–183. https://doi.org/10.7901/2169-3358-2014.1.170

Nylund, M. (2022). Community-Based Participatory Research. Springer EBooks, 1–3. https://doi.org/10.1007/978-3-319-69909-7_486-2

O'Connor, R. C., Wetherall, K., Cleare, S., McClelland, H., Melson, A. J., Niedzwiedz, C. L., O'Carroll, R. E., O'Connor, D. B., Platt, S., Scowcroft, E., Watson, B., Zortea, T., Ferguson, E., & Robb, K. A. (2020). Mental health and wellbeing during the COVID-19 pandemic: longitudinal analyses of adults in the UK COVID-19 Mental Health & Wellbeing study. The British Journal of Psychiatry, 218(6), 1–17. https://doi.org/10.1192/bjp.2020.212

Online Form Builder | Create Online Forms for Free – Zoho Forms. (n.d.). Zoho. https://www.zoho.com/forms/

Palafox, B., Renedo, A., Lasco, G., Palileo-Villanueva, L., Balabanova, D., & McKee, M. (2020). Maintaining population health in low- and middle-income countries during the COVID-19 pandemic: Why we should be investing in Community Health Workers. Tropical Medicine & International Health, 26(1), 20–22. https://doi.org/10.1111/tmi.13498

Park, J. (2021). Who is hardest hit by a pandemic? Racial disparities in COVID-19 hardship in the U.S. International Journal of Urban Sciences, 25(2), 1–29. https://doi.org/10.1080/12265934.2021.1877566

Pinto, D., Carroll-Scott, A., Christmas, T., Heidig, M., & Turchi, R. (2020). Community health workers: improving population health through integration into healthcare systems. Current Opinion in Pediatrics, 32(5), 674–682. https://doi.org/10.1097/mop.0000000000000940

Porat, T., Nyrup, R., Calvo, R. A., Paudyal, P., & Ford, E. (2020). Public Health and Risk Communication During COVID-19—Enhancing Psychological Needs to Promote Sustainable Behavior Change. Frontiers in Public Health, 8. https://doi.org/10.3389/fpubh.2020.573397

Rao, R., Plotnick, L., & Hiltz, R. (2017). Supporting the Use of Social Media by Emergency Managers: Software Tools to Overcome Information Overload. http://hdl.handle.net/10125/41185

Rathore, V. (2016). Technology in Disaster Management and Disaster Risk Reduction: A Review of Applications. https://core.ac.uk/download/pdf/234664575.pdf

Russell, D., Oberlink, M. R., Shah, S., Evans, L., & Bassuk, K. (2018). Addressing the Health and Wellness Needs of Vulnerable Rockaway Residents in the Wake of Hurricane Sandy. Journal of Public Health Management and Practice, 24(2), 137–145. https://doi.org/10.1097/phh.0000000000000545

Sakurai, M., & Murayama, Y. (2019a). Information Technologies and disaster management – benefits and issues -. Progress in Disaster Science, 2, 100012. https://doi.org/10.1016/j.pdisas.2019.100012

Scheib, H., & M. Brinton Lykes. (2013). African American and Latina community health workers engage PhotoPAR as a resource in a post-disaster context: Katrina at 5 years. Journal of Health Psychology, 18(8), 1069–1084. https://doi.org/10.1177/1359105312470127

Stevenson, M. C., Norrbom, C., Savela, M., Xiong, Y. L., Lee, T. F., Garcia, C., Winstead, O., Northrop, M., & Sandy, M. (2022). Community Health Workers in time of crisis: A covid-19 case study. Journal of Humanistic Psychology, 002216782211327. https://doi.org/10.1177/00221678221132718

Shklovski, I., Palen, L., & Sutton, J. (2008). Finding community through information and communication technology in disaster response. Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, 127–136. https://doi.org/10.1145/1460563.1460584

Tai, D. B. G., Sia, I. G., Doubeni, C. A., & Wieland, M. L. (2021). Disproportionate Impact of COVID-19 on Racial and Ethnic Minority Groups in the United States: a 2021 Update. Journal of Racial and Ethnic Health Disparities, 9(6), 1–6. https://doi.org/10.1007/s40615-021-01170-w

Taylor, B., Mathers, J., & Parry, J. (2017, April 7). Who are community health workers and what do they do? development of an empirically derived reporting taxonomy. OUP Academic. https://academic.oup.com/jpubhealth/article/40/1/199/3111232

The COVID Tracking Project. (2021). Wisconsin: Race & Ethnicity Historical Data. [online] Available at: https://covidtracking.com/data/state/wisconsin/race-ethnicity/historical [Accessed 26 Dec. 2023].

U.S. Department of Health & Human Services. (2019). Role of Community Health Workers, NHLBI, NIH. Nih.gov. https://www.nhlbi.nih.gov/health/educational/healthdisp/role-of-community-health-workers.htm

USAFacts. (2023). Milwaukee County, Wisconsin coronavirus cases and deaths. [online] Available at: https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/state/wisconsin/county/milwaukee-county/ [Accessed 26 Dec. 2023].

Wennerstrom, A., Vannoy, S., Allen, C., Meyers, D., O'toole, E., Wells, K., Springgate, B., & Dis, E. (2013). Community-Based Participatory Development of a Community Health Worker Mental Health Outreach Role to Extend Collaborative Care in Post-Katrina New Orleans NIH Public Access Author Manuscript.https://chwcentral.org/wp-content/uploads/2013/11/Wennerstrom_CB-Participatory-Development-of-CHW-Mental-Health.pdf

Wisconsin Department of Health Services. (2020). COVID-19: Wisconsin Summary Data. [online] Available at: https://www.dhs.wisconsin.gov/covid-19/data.htm.

World Health Organization. (2020, January 26). Risk communication and community engagement (RCCE) readiness and response to the 2019 novel coronavirus (2019-nCoV). Www.who.int. https://www.who.int/publications/i/item/risk-communication-and-community-engagement-readiness-and-initial-response-for-novel-coronaviruses

Psychosocial Aspects of the 2021 Flood Disaster in Euskirchen District, Germany: Strains, Resources, and the Role of Social Media

Francesca Müller*
University Wuppertal
framueller@uni-wuppertal.de

Marvin Kubitza
University Wuppertal
marvin.kubitza-hk@uni-wuppertal.de

Frank Fiedrich
University Wuppertal
fiedrich@uni-wuppertal.de

ABSTRACT

Purpose: With the increasing frequency of disasters, the focus of this paper is on the acute and long-term stress and strain resulting from the 2021 flood disaster and the role of social media in this. Methods: Data was collected through group discussions in the context of world cafe´ sessions. In April 2023, three workshops were carried out with various groups of people in the Euskirchen district. Minutes and transcripts were analysed using Mayring's qualitative content analysis. Results: The results indicate a high level of stress for those affected due to environmental, social and psychological factors on an analogue and digital dimension. The strain was usually recognised and acknowledged in the aftermath of the crisis situation. Conclusions: It is recommended that the inhibition threshold for mental health and psychosocial support be lowered further and that a psychosocial situation report be kept over a significantly longer period of time.

Keywords

natural disasters, stress and strain, social media, disaster management, mental health and psychosocial support

INTRODUCTION

In recent years, combined hazard situations in the form of geophysical, hydrological and meteorological crises and disasters (CaD) have increased (UNDRR & CRED, 2020). In particular, the number of natural disasters, especially floods, has almost doubled in the last 40 years (ibid.). The significant impact of natural disasters on a large proportion of the population affects both people's lives and the environment (Berz, 2014). CaD are characterised by a sudden and unexpected discrepancy between perception and coping, as well as heterogeneous and significant impacts on the well-being and functioning of those affected (Quarantelli, 2005). A width of sidewalk can separate a fate of destroyed livelihood from integrity. Those affected show individual reactions and psychological consequences, as well as subjective perceptions of the circumstances. In particular, the emotional and psychological aspects associated with such events, such as fear, loss and survival instincts, depend on individual experiences and characteristics. Mental health is a crucial aspect of overall preparedness and crisis management. Mental health and psychosocial support (MHPSS) aims to reduce the likelihood of clinical symptoms occurring in the aftermath of a CaD. Even today, in 2023, many counselling centres are still maintained and accepted by the population.
In CaD, the available coping resources and basic environmental, psychological and social needs change. As a result, those affected have an increased need for information and communication. Because of their interactive nature, social media (SoMe) provide a platform for private and public exchange. SoMe are already part of everyday life for 70 % of

*corresponding author

users (Paulsen et al., 2023). Among other things, people can share their thoughts, feelings and challenges with and receive support from others (Grieve et al., 2013). As a result, SoMe is increasingly being used in CaD to manage the crisis situation (Norris et al., 2008). To date, scientific research has focused thematically on the needs in the acute phase of CaD and, in terms of target groups, primarily on authorities and emergency personnel (AEP). This article therefore focuses on stress and resource-strengthening factors in the acute and reconstruction phase (RQ1) as well as perceptions and experiences with MHPSS services (RQ2). Starting with a general description of the 2021 flood disaster, scientific findings on stress factors and resources in crisis situations and MHPSS, a detailed description of the methodology used and the data basis is provided. The derivation of the research questions is integrated into the description of the associated studies. Taking these principles into account, the results are presented, discussed comparatively and methodological limitations are highlighted. Finally, specific recommendations for decision makers and conclusions are drawn.

BACKGROUND AND RELATED STUDIES

2021 European Floods

Natural disasters are extreme events caused by natural phenomena that have a variety of characteristics, including the type of underlying hazard, speed of occurrence, duration, predictability, warning potential and magnitude of impact. They can be categorised into geophysical, hydrometeorological and biological disasters, with floods, earthquakes and hurricanes being the most common (Jha, 2010).

The flood disaster that occurred in many European countries in July 2021 was a devastating natural disaster caused by heavy and prolonged rainfall.The countries affected included Germany, the United Kingdom, the Netherlands and others.Caused by a low-pressure system called "Bernd", intense rainfall occurred within a few days, causing water levels to rise rapidly and causing many rivers to burst their banks. The masses of water led to landslides and the destruction of settlements and infrastructure (Jungha¨nel et al., 2021). Around 180 German cities and communities were affected by the flooding (Radio Euskirchen, 2022). The federal states of North Rhine-Westphalia (Euskirchen district and Hagen) and Rhineland-Palatinate (Ahrweiler and Bad Neuenahr-Ahrweiler) were particularly affected. In some areas, 150 litres/m² were measured within 24 hours (compared to the average rainfall in Germany in 2021 of 805 litres/m² (Deutscher Wetterdienst, 2023)). The flood disaster lasted several days. Although meteorologists had warned of the intense rainfall, the extent and spatial distribution of the rainfall was difficult to predict. In total, more than 180 people died in Germany, about 48 of them in North Rhine-Westphalia (Schro¨der, 2022).

Stress Factors and Resources in Crises and Disaster situations

Every CaD is a massive collective stressor for the population (Vitaliano et al., 1987), which affects individuals in different ways and forms at the micro (individual), meso (organizational) and macro (societal) level. Personality as well as temporary physical and psychological states act as moderators (Paton & Flin, 1999). Consequently, there are significant individual differences in response to emergency stressors that influence the magnitude of the perceived stress response. The term *stress* refers to all influences affecting a person, both positively and negatively, without any subjective evaluation. Depending on the type of CaD the incurred stress and its processing vary (Galea et al., 2005). Strain can be understood as a negative emotional experience aimed at reactive adaptation in the form of predictable biochemical, physiological and behavioral changes. According to Baum, a distinction can be made between changing the external stimulus through manipulation of the situation and changing the internal effect as an adjustment to the individual interpretation (Baum, 1990). The stresses that AEPs and the population can face have been extensively researched (Leon, 2004) and the physical and psychological health impacts of natural disasters have also been highlighted (Makwana, 2019).

Nevertheless, the perceptions and external stimuli are changing due to the paradigm shift from a passive population to prosumers (Lang et al., 2020). Those affected communicate through SoMe to actively share and search for information (Reuter & Kaufhold, 2017)This can also give rise to stressors from the digital space, such as disruptions in people's lives, relationships and communities (Carroll et al., 2009). Among other things, long periods without new information, long working hours, lack of social support, noise and hazardous substances as physical/chemical factors and reduced boundaries between work and private life can have psychological impact. In summary, there is a need for a comprehensive understanding of stress that takes environmental (primary), psychological and social (secondary) stressors into account (Pearlin, 1989). The factors cannot be viewed in a clearly differentiated manner, but rather determine each other in causal relationships.

The **environmental impacts** of flood disasters include impacts on water quality and infrastructure, as well as the situational factors such as noise and stale smell. Environmental conditions can also have an impact on psychological distress. Noise can lead to hearing damage and trigger stress reactions with long-term exposure to 70–80 dB(A) (Gross, 2012). **Social stressors** in flood disasters are diverse and include both immediate and long-term impacts

in interaction with the social environment, such as significant life changes in relationships or socioeconomic disadvantage (Rufat et al., 2015). In the context of crisis communication, it is already known that defensive crisis communication can damage the credibility of individuals or authorities (Allga¨uer & Larisch, 2011) and that it is better to communicate negative information openly and transparently than to hide it (Frewer et al., 2002). Affected people actively search for information, including on SoMe. Furthermore, they disseminate information rapidly and in some cases provide specific situational information that is not available through traditional media (Liu & Austin, 2010). Exposure to flooding can have significant **psychological impacts**, with symptoms of post-traumatic stress disorder, depression or anxiety (Mason et al., 2010). These effects can be compounded by secondary stressors such as health concerns, relationship problems (social), and loss of sentimental items (environmental) (Tempest et al., 2017). In addition, encountering the suffering of others can lead to secondary trauma (Daniels, 2008). Due to the user-generated nature of SoMe, an exchange of information can occur that can lead to exposure to traumatizing media content among unaffected populations (Kreutzer & Köllner, 2020) or to rumors (Nestler, 2017). SoMe has become not only a stressor but also a crucial resource in CaD, fulfilling various functions (Mu¨ ller et al., 2023). These primarily include social factors such as information dissemination, resource mobilization and psychosocial support through emotional relief and strengthening a sense of community (Simon et al., 2015)Because psychological strain is not only a consequence of the stressors, but also a result of the way a person perceives and evaluates the situation, taking into account their own resources (Lazarus & Folkman, 1984), there is still a lack of knowledge about the long-term strains and resources of the population during CaD (Walker-Springett et al., 2017). Perception is influenced by various factors. One factor is anchoring bias, which illustrates how individuals use a given starting point or anchor when evaluating information or values and anchor their evaluations to that reference point (Tversky & Kahneman, 1974). Another factor, especially in stress-dominated situations, is selective perception. This is characterized by the rapid interpretation of information through the individual lens of one's own expectations, perspectives and experiences, despite the belief that reality is perceived objectively (Glaser, 2019). Furthermore, the negativity bias causes the tendency to pay much more attention to negative information than to positive information (Vaish et al., 2008). In particular, there is a need for further research to examine today´s stresses and resources of the German population after a specific event. Therefore, we propose the first research question (RQ1).

RQ1: Which stress and resource-strengthening factors were particularly relevant for those affected by the 2021 flood disaster in the acute and reconstruction phase?

As SoMe play a crucial role in today's CaD, we also focused on stresses and resources in the digital space.

Mental Health and Psychosocial Support in Crises and Disaster Situations

Since the success of physical reconstruction after CaD also depends on the psychological functioning of the population, early interventions to prevent psychological disorders after natural disasters are particularly important (Kessler et al., 2012). In CaD MHPSS as a system of psychosocial interventions includes (1) strengthening individual and social coping resources, (2) promoting resilience and (3) preventing psychological trauma disorders (Beerlage, 2021). The measures relate to CaD and connected stressful operational situations (e.g. death or danger to one's own life). Grawe distinguishes four psychosocial basic needs: (1) attachment, (2) orientation and control, (3) enhancement of self-esteem, and (4) pleasure gain/avoidance of displeasure. Attachment is the need for interpersonal relationships. Social support therefore plays an important role in individual psychology and has a positive effect on stress management, adaptability and life satisfaction (Cohen & Wills, 1985). Conversely, the lack of social connections can have negative effects on mental health (Baumeister & Leary, 1995). The return to one's own ability to act as an essential act of psychosocial support promotes the satisfaction of orientation and control needs as well as the strengthening of self-esteem (Kernis, 2003).

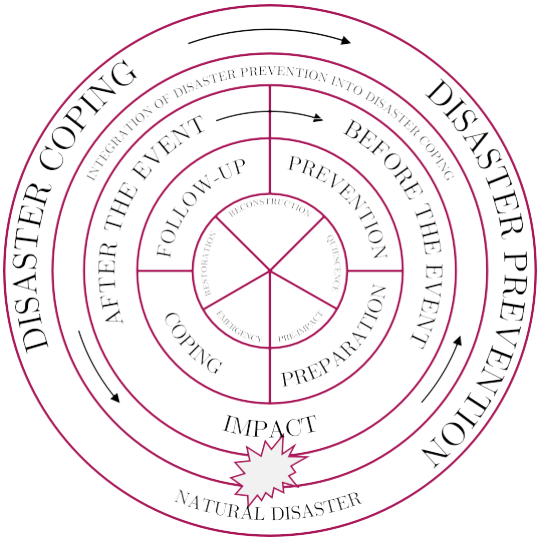


Figure 1. Disaster cycle, combined and revised from (BBK, n.d.; Glade & Alexander, 2013)

MHPSS measures vary depending on the individual characteristics and reactions. However, they also vary depending on the current phase of the disaster management cycle. Figure 1 visualizes this cycle at different levels based on

(Glade & Alexander, 2013). In principle, crisis management is divided into the four cyclical phases of 1. prevention, 2. preparation, 3. coping and 4. follow-up (BBK, n.d.). The psychosocial measures are divided into A pre-activity phase, B activity phase and C the phase after the disaster response (Umeda et al., 2020), whereby phase C can be further broken down on a time scale of two months to several years into relief phase, rehabilitation phase and rebuilding phase (Rao, 2006). Phase A is particularly characterized by preparatory measures aimed at experiencing stressful situations and one's own perception of stress. Due to their preventative nature, these measures are primarily carried out for AEPs (Umeda et al., 2020). The activity phase (Phase B) is characterized above all by immediate physical care and medical care, but also by connecting survivors with sources of support, culturally appropriate forms of mourning for the deceased and normalization of activity and routine as far as possible (Rao, 2006). After the disaster response, activities are geared towards returning to routine and sharing and processing experiences (Umeda et al., 2020). Here (relief phase), primarily empathic and active listening as well as relaxing methods are necessary to support those affected (Rao, 2006). The rehabilitation phase is characterized by the realization of those affected that they must become active themselves and help themselves. Family support and the social environment as well as acceptance and supportive reactions from others are particularly important here (ibid.). Detached from the disaster cycle, MHPSS can be differentiated into two dimensions: (A) temporal and (B) target group-oriented. MHPSS actors act in the short, medium and long-term support, advice and therapy for those affected in CaD (A) (Beerlage, 2021). The target group-oriented dimension (B) takes certain groups of affected people into account when assessing the need for MHPSS measures due to different initial conditions, needs and perspectives. A distinction is made between B.1 measures for survivors, relatives, witnesses and/or missing persons and B.2 measures for AEP. Measures for affected group B.1 are primarily aimed at satisfying the needs for security, reassurance, regaining self-efficacy/control, connection with loved ones, confidence and information (psychoeducation) (Mu¨ ller-Cyran & Zehentner, 2013). Measures for affected group B.2, on the other hand, also focus very much on preventative measures as part of the integration into health and occupational safety and the employer's duty of care (Beerlage, 2021).

Due to the individuality of reactions and coping processes, it is not only important to understand the perceptions and experiences of those affected and to gain insights into the effectiveness of MHPSS interventions, but also to adapt future psychosocial support services appropriately to the needs of the community. MHPSS interventions are a developing field with few established interventions. This in turn limits the level of awareness of available interventions, as many of those affected are unaware of their existence and implementation. Due to the sensitive nature of personal data, there are numerous data protection restrictions when contacting people who are being supported. An open research question in this thematic field is therefore the RQ2 considered here.

RQ2: How did those affected by the 2021 flood perceive and experience the MHPSS services?

METHOD

As part of the study, workshops were held with people affected by the flood disaster in 2021 in the Euskirchen district, Germany. Qualitative research makes it possible to survey opinions, motives and attitudes in a rule-based manner.The workshops took place on three different occasions with different groups of participants who were involved in the disaster in different ways, see table1. During the workshops, the groups engaged in a dialog using the "world cafe´" technique. The debates were documented for each topic and then transcribed. The transcripts were analyzed in accordance with the research questions using qualitative content analysis according to (Mayring, 1991). Figure 2 visualizes the basic process from recruiting interviewees to analysis.

Participant Recruiting and Data Collection

In cooperation with the Euskirchen district, a large-scale open web survey was carried out in March 2023. Integrated into this survey was a request for further willingness to participate in other exchange formats and consequently the recruitment of participants for the group discussions. The survey was disseminated by the University Wuppertal and the Euskirchen district via various communication channels (radio, SoMe, etc.). As a result, over 1,400 complete sets of responses were obtained from those affected by the 2021 flood disaster as well as spontaneous helpers and AEPs. 215 persons were willing to participate in group discussions (101 of them with characteristics of the target groups). Based on specific queries, three group discussions were held, each with at least 12 persons from different target groups. The discrepancy in the number of people between those on standby and participants is due to logistical restrictions, lack of acceptances and no-shows. Within two weeks in April 2023, thus almost 2 years after the flood disaster 2021, different topics were examined from different perspectives on three different days, see table 1.

The methodical implementation took the form of a world cafe´, a process that records discussions on various issues in rounds and with changing group composition and thus offers a supplement to other established methods of qualitative data collection (Lo¨hr et al., 2020). As a participatory method, this form of group discussion not only

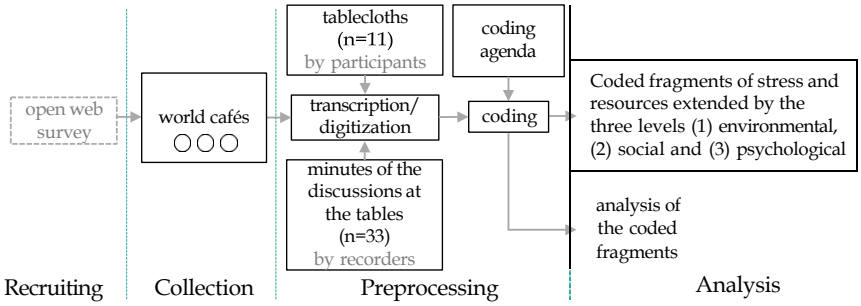


Figure 2. Framework methodology process

provides data for the researchers, but can also be useful for the participants, as they can reprocess their experiences from the 2021 flood disaster in dialog and psychoeducationally facilitate the recording of their own needs (Steier et al., 2008). The participants' discussions were led by the hosts and recorded in writing by minute-takers. In addition, the participants were also asked several times to note down important key points and topics on writable tablecloths. All hosts of the three simultaneous discussion rounds were instructed in advance by the project management and prepared for the methodology with the help of rules and sub-questions on the discussion topic. Minutes were taken using a semi-structured template.General notes on the course of the discussion were recorded, and frequent or controversial topics were explicitly noted. In order to be able to identify particular focal points across the different groups, the note-takers were also asked to note topics discussed twice and to set time stamps.

Table 1. Topics of the world cafe´ sessions

date 1 *AEPs*	**date 2** *affected persons who have made use of MHPSS*	**date 3** *helpers of MHPSS*
special stress and resilience factors	special stress and resilience factors	classification of requirements for MHPSS
use of SoMe		
preparation and training	development wishes MHPSS	development wishes MHPSS
role of SoMe		

Data Preprocessing and Analysis

Both the protocols and tablecloths were transcribed after the events and coded deductively-inductively as part of a content-structuring qualitative content analysis (Schreier, 2014). The core of the content-structuring approach was to identify and conceptualize categories within the framework of selected content aspects and to systematically describe the material with regard to stressors, resilience factors and experiences as well as perspectives on MHPSS. Based on the modified version of the procedure according to Steigleder, a combined deductive-inductive formation of super- and subcategories was implemented. The upper categories were formed thematically, while the subcategories were identified as in vivo codes (Steigleder, 2008). A total of N=996 codes were allocated to a category, with the possibility of multiple codings.

RESEARCH FINDINGS

The 2021 flood disaster (14/15.07.2021) was a natural disaster with severe flash floods and various damages. 73 % of respondents to our open web survey indicated at least a medium level of psychological stress in relation to the impact on their own household (approx. 10 % no stress at all), more than 87 % in relation to the wider social environment. Up to one month after the flood disaster, the subjectively perceived at least medium high stress is assessed by the majority (81 %) due to the impact on the immediate social environment (4 % not at all, 5 % very low, 10 % low), approx. 67 % in relation to the impact on their own household (15 % not at all, 9 % very low, 9 % low). Overall, both primary and secondary traumatization put a heavy strain on the population, which extends beyond the coping phase.

Stress and Resource-strengthening Factors During the Flood Disaster 2021 (RQ1)

During the emergency phase (see figure 1), the focus of those affected and AEPs was on saving lives and taking the acutely necessary measures to prevent or minimize damage. During the group discussions, it became clear that the focus of all participants with different degrees of involvement was and is on the stress factors and less on the resilience factors. In a direct comparison, 57 % of the coded fragments were categorized in the "stressors" category, 43 % in the "resources" category (n_{total} = 402 coded fragments). Risk and crisis communication in relation to the flood disaster was discussed very controversially. The information required to assess the risk was not received by everyone and was also interpreted differently. In particular, the indication of the expected amount of water in l/m² was negatively emphasized, as many did not know how to interpret the warning. Crisis communication was felt to be inadequate not only before the prevention phase, but also during the disaster coping phase. The lack of communication about what was happening in the Euskirchen district led to uncertainty. Furthermore, the extent of the damage in the Euskirchen district was hardly discussed during the coping phase. This led to a feeling of not being seen and not having one's suffering acknowledged until the group discussions almost two years after the event. However, the psychosocial stress was overshadowed by the material and acute life-threatening stress during the coping phase. Figure 3 schematically illustrates this temporal relative relationship. In order to coordinate crisis management and to react appropriately to a CaD a general situation report is created and maintained. The psychosocial situation report, as a specialised form, supplements the overall picture with the aspect of the psychological factors and needs of the population groups. The need for MHPSS depends on various factors and the course of mental coping is described as fluctuating, but above all as lasting. Consequently, the psychosocial situation reportand crisis management must be maintained beyond the period of the general situation report. During the group discussions, noise and specific smells, such as the musty smell of the basement, were still described as significant stress factors in the coping phase and triggers for physical and emotional reactions, even two years after the flood disaster. In the first few weeks after the event, there was no way to mitigate the noise pollution through cleanup, heavy machinery and other vehicles, or to find a quiet place to rest. The sludge and exhaust gases emitted by the machines led to unpleasant odors in the long term. In addition to these factors, rumors in particular were mentioned as a burden from the digital space (especially SoMe) with an impact on analogue events. What is particularly striking here is that it was not the quantitative amount of rumors, but rather the impact of individual ones that was actively discussed by all groups. A key point here were the so-called "flood tourists". The resulting burden was not necessarily due to the value of the looted items, but rather to the fact that those affected could not imagine looters wanting to profit from their suffering. The same effect was achieved by people who traveled to the flood areas for self-promotion purposes on SoMe in order to supposedly help there and thus generate more followers.

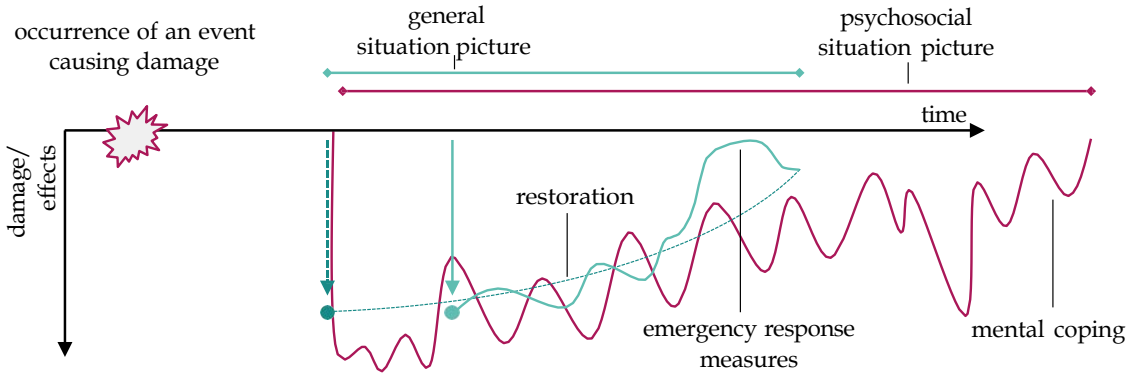


Figure 3. Schematic progression of crisis management, broken down by material/environmental (mint) and psychological consequences (pink), revised and expanded from (BSI, 2023)

SoMe were highly valued as a communication medium. First, it allowed individuals to reach out to friends and family and inquire about their well-being. Second, it made it easier to organize self-help groups and relief campaigns. SoMe was used to launch appeals and set up support groups. After the disaster, thousands of volunteers from all over Germany traveled in cars and buses to help. Donations of relief supplies were also received from large corporations. In addition, SoMe played a role in informing the public during the disaster. However, those affected noted that communication was not always possible. During the acute phase and shortly afterwards, the mobile phone network collapsed, making telephone calls and internet access impossible. The inability to monitor the well-being of their social environment and the lack of protection from external factors were identified as sources of stress. SoMe were therefore, only when usable, a resource to strengthen the community. The sense of community between

those affected and those helping them was cited several times by all groups as the most important resource-creating factor. The gratitude for the large number of people who selflessly helped in the CaD areas was still discussed immensely two years after the event.

Perception and Experiences with the MHPSS Services (RQ2)

The perception of circumstances is subject to many influences that can distort them. As a result, even the same stresses lead to different strains. The significantly different levels of stress in the course of the flood disaster led to a population with very different levels of strain. During the group discussions, many of the participants rated their own stress in relation to the known extent of damage to others. This also resulted in a corresponding classification of their strain, which in turn led to a trivialization of their reaction. On the micro level (individual), this showed an alignment of the reported stress with the macro level (group-specific), which was also visible between the different dates and target groups. Both those affected and the AEPs described communication problems, noise, smells and flood tourists as stress factors. However, it was also often stated that the stress on others from the immediate or wider social environment caused individual reactions, even to the point of strain. A danger from this can be that as the comparison number increases with SoMe, the subjective feeling of strain and the probability of seeking help decreases.

One significant difference between the target groups of AEP and those affected was in the perception of MHPSS. While both groups rated the services extremely positively, for the AEP the service was clearly known and sometimes mandatory in group settings, while not all those affected had information about the services (in the open web survey, almost 30 % of respondents said MHPSS was not known). The timing for disseminating information from MHPSS services was explicitly highlighted in the discussions. In the early stages of the coping phase, several initiatives, brochures and MHPSS services were provided. The offered MHPSS services were available during the day, particularly during clean-up operations. Some affected people said they were unable to accept such offers because of their focus on the cleanup efforts. From the perspective of the helpers of MHPSS (third date), the demand for MHPSS services exceeded the availability of various MHPSS offerings around three months after the event. The use of support services was perceived to be very fluctuating. Coping with a CaD presents individuals with numerous challenges in terms of preparation, assessing the situation, making decisions and dealing with stress (Enander, 2021). This process is not linear, see also figure 3. Although many people are open to using the support services, the fluctuating psychosocial demands and the partial lack of self-reflection result in the intensity of the population's use of MHPSS services fluctuating over time. Building on this, it was repeatedly pointed out in all three groups that the inhibition threshold for MHPSS measures needed to be lowered and awareness of the need for such support needed to be increased. The fear of social repercussions such as disrespect from one's own environment or the stress of having to explain oneself often outweighs the inclination to seek help. The terminology "psyche" also plays a negative role in the evaluation. In order to lower the inhibition threshold, the participants suggested measures such as active contact, more advertising, transparent information or organized meetings in cafe´s or similar places. A combination of active and passive (come and go structure) should be aimed for here.

DISCUSSION

The aim of the study is to investigate stress and resource-inducing factors of the 2021 flood disaster, with a special focus on the digital space (SoMe) and experiences with MHPSS services. Three group discussions were conducted as world cafe´s with affected persons, AEP and MHPSS staff in the Euskirchen district. Participants mainly focused on stress-triggering factors, and the psychosocial situation outlasted the general crisis. The digital situation significantly influences the analogue situation (RQ1). MHPSS, as a resource-enhancing factor, was perceived positively in the group discussions, but awareness was rated as insufficient and utilisation was not low-threshold enough (RQ2).

Comparative Discussion

The results of the structuring content analysis align with other scientific findings. The **focus on negative information**, encompassing stress-inducing factors, has been previously discussed under the term negativity bias (Vaish et al., 2008). The identified referencing of one's own stress to the macro level has similarly been analysed in the context of the anchor effect and selective perception (Glaser, 2019; Tversky & Kahneman, 1974). The anchor serves as a comparative value referencing a negative experience of another person in the context of the anchor bias. In the context of selective perception, those affected assume that others are more affected and neglect their own affectedness. To date, there is no external evidence that individuals downplay their own impact in CaD. Similarly, there is a lack of evidence as to whether others underestimate the extent to which they are affected by such disasters or whether this

influences the utilisation of MHPSS. Nevertheless, Plapp (2003) examines this issue and concludes that personal relationships with the affected area determine the degree to which CaD are relativised. Further research in this area is necessary to analyse the identified "**diffusion of stress**" more deeply. Further, **crisis communication** should be comprehensible, precise, specific, call for action and locally controlled (Aguirre, B. E., 2014). Communication should take place via various media, including SoMe. The provision of information via SoMe was a key factor for those affected. The lack of information and the lack of attention to the extent of the damage in the media was perceived as an additional stressor.

Frewer et al. (2002) argues that some information, even if insufficient, is preferable to no information at all. An information status display also reduces the likelihood of rumors, frequently discussed as a negative factor in the digital space affecting analog events. However, it's not the quantity of misinformation that matters, but rather its quality, credibility, and dissemination. The study by Fathi and Fiedrich (2022) revealed that only 5 % of social media posts related to the flood disaster were misinformation, yet this could have a significant impact.
Exposure to **noise and odour** is particularly relevant in the area of occupational health and safety. Loud and persistent noise poses a potential risk to health and is categorised as a hazard in the occupational health and safety risk assessment. The multifactorial exposure causes a link between the environmental conditions and the emotions and feelings. This link between sounds or images and the triggering of post-traumatic stress disorder has already been well researched (Pausch & Matten, 2018). Interestingly, the sight, i.e. physically seeing the flooding and destruction, was not explicitly mentioned as a stress factor in the group discussions conducted. Further research is needed to determine whether this was simply not addressed due to its obviousness or whether it may be a case of desensitization from watching disasters on TV or similar.
The shared experiences of disaster-affected individuals foster the development of **group affiliation** or community, recognized as a crucial factor in resource mobilization. According to the social identity theory proposed by Tajfel and Turner, people tend to categorize themselves and others based on shared characteristics like gender, ethnicity, and religion, leading to community formation. Social media (SoMe) facilitated communication within and beyond these communities, a point underscored during discussions. Currently, organizational and communication practices via SoMe are considered typical Zisgen et al.

The coping response in CaD may require MHPSS services, yet the reluctance to seek counseling remains a significant issue. While some guidelines acknowledge the challenge of seeking psychological help, the existing barrier is often deemed too high (Nikendei, 2012). To address this, the group discussion suggested promoting active contacts and transparent information provision by the Federal Office of Civil Protection and Disaster Assistance. The study's results indicate that psychological stress lingers long after the event, suggesting a prolonged need for psychosocial situational awareness. Solomon and Green suggests psychological consequences can last up to three years, with most symptoms subsiding after 16 months (Solomon & Green, 1992). The peak of symptoms occurs one year after the disaster, followed by gradual improvement (Goldmann & Galea, 2014).

Methodological Limitations

The chosen combinatorial structuring content analysis methodology, relying on minutes and tablecloths from various world cafe´s, has limitations. Participants were recruited through an open web survey, lacking random sampling or representativeness. This generated a sample of intrinsically motivated individuals willing to participate in both the survey and group discussions. The extent of the damage made it difficult to clearly distinguish between the groups of people on the three dates. The group formation and invitation was carried out in waves, whereby acceptances from the participants' social environment were also taken into account. The study aimed for qualitative insights, not quantitative data. Gathering information from other regions could enhance findings. Conducted almost two years post-flood, there's a risk of recall bias alongside selection bias. Distance to the event may distort memory accuracy. However, selective perception in CaD doesn't weaken surveyed factors' relevance. Methodological limitations include information bias due to different hosts and minute-takers impacting participant involvement. For reasons of clarity, no distinction was made between stress and strain in the group discussions, meaning that it was no longer possible to differentiate between them when coding. In addition, the content analysis lacked a reliability check. Further validation is planned through open web survey and interview analyses, with knowledge application in practice through practitioner exchanges.

Implications for Decision Makers

Implications for decision makers in CaD, drawn from group discussions with affected persons, AEPs, and MHPSS staff of the 2021 flood disaster, include:

1. Establish trust in SoMe early in the prevention phase for effective risk and crisis communication. Ensure open communication, sufficient personnel, and clear language. Use practical comparisons and explicitly name potential hazards to raise awareness and assess individual vulnerability. Consider that naming the extent of damage is crucial for those affected.
2. Stress from rumors in the digital space is influenced by proximity and reach, not volume. Early denial by public bodies, supported by credible SoMe accounts and consistent communication, can mitigate this stress. Enhancing population media skills is essential.
3. Support the population during the restoration phase by distributing personal protective equipment, such as earplugs or odor inhibitors.
4. Increase awareness of psychosocial stress and MHPSS services. Psychoeducational measures should cover stress reactions, secondary traumatisation, and the use of MHPSS. Offer services in a flexible structure integrated into community activities during the coping phase.
5. Maintain a long-term psychosocial situation picture and crisis management beyond the coping phase.

CONCLUSION

The flood disaster in July 2021 caused deaths and extensive damage due to heavy and prolonged rainfall, affecting people across regions. They faced challenges such as unclear crisis communication, theft, and the loss of loved ones. The changing society and the rise of digital communication led to the first research question:
RQ1: What stress and resource-strengthening factors were particularly relevant for those affected by the 2021 flood disaster in the acute and reconstruction phase?

Group discussions highlighted analog stress factors, like noise and stale smell, with a strong emphasis on the lack of crisis communication, including via SoMe. Various groupings noted that information from SoMe impacted physical events and actions of those affected. Discussions suggested addressing population needs not in the acute phase but during coping. The interrupted flow of communication and a lack of information about the social environment, due to everyday use, added an additional burden, deviating from everyday life in today's networked world.

MHPSS played a major role in processing deviations from everyday life in the 2021 flood disaster. To gain further insights into CaD with a focus on MHPSS, the study explored the second research question:
RQ2: How did those affected by the 2021 flood perceive and experience the MHPSS services?

While MHPSS services are not universally known, group discussions predominantly reported positive experiences. Various development wishes were discussed, indicating optimization potential in this area.

Further research is crucial due to the increasing likelihood of natural disasters. Investigation into factors contributing to stress and resources in CaD in today's society is essential. Identifying consistent factors through comparisons with other disasters and assessing the effectiveness and acceptability of MHPSS in CaD require additional research. As part of the research project #sosmap, further quantitative and qualitative data will be collected to shed light on the current status of the MHPSS and explicitly the role of SoMe from the various perspectives of the population as well as the authorities and emergency services. The data already collected as part of the study (survey results and interview transcripts) will be analyzed in greater depth with various objectives in mind. These include the perception of the MHPSS and a comparison between self-reported use of SoMe compared to identifiable information in publicly available sources.

REFERENCES

Aguirre, B. E. (2014). *Planning, warning, evacuation, and search and rescue: A review of the social science research literature*.

Allga¨uer, J. E., & Larisch, M. (Eds.). (2011). *Public Relations von Finanzorganisationen: Ein Praxishandbuch fu¨ r die externe und interne Kommunikation*. Gabler Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden. https://doi.org/10.1007/978-3-8349-6770-1

Baum, A. (1990). Stress, intrusive imagery, and chronic distress. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, *9*(6), 653–675. https://doi.org/10.1037/0278-6133.9.6.653

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin*, *117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Beerlage, I. (2021). Psychosoziale Notfallversorgung (PSNV). https://doi.org/10.17623/BZGA:Q4-I140-1.0

Berz, G. (2014). Risiken und Chancen aus Naturkatastrophen und Klimawandel: Geowissenschaftliche und (versicherungs-) wirtschaftliche Perspektiven. In H.-G. Bohle & Pohl, Ju¨rgen, Blu¨mel, Wollf Dieter (Eds.), *Risiko: Erkundungen an den grenzen des wissens* (pp. 79–89). Wissenschaftliche Verlagsgesellschaft Stuttgart.

Bundesamt fu¨r Bevo¨lkerungsschutz und Katastrophenhilfe. (n.d.). Krisenmanagement ist ein Zyklus.

Bundesamt fu¨r Sicherheit in der Informationstechnik. (2023, June). *Business continuity management: BSI-Standard 200-4*.

Carroll, B., Morbey, H., Balogh, R., & Araoz, G. (2009). Flooded homes, broken bonds, the meaning of home, psychological processes and their impact on psychological health in a disaster. *Health & place*, *15*(2), 540–547. https://doi.org/10.1016/j.healthplace.2008.08.009

Cohen, S., & Wills, T. A. (1985). *Stress, social support, and the buffering hypothesis* (Vol. 98). https://doi.org/10.1037/0033-2909.98.2.310

Daniels, J. (2008). Sekunda¨re Traumatisierung: Interviewstudie zu berufsbedingten Belastungen von Therapeuten. *Psychotherapeut*, *53*(2), 100–107. https://doi.org/10.1007/s00278-008-0585-y

Deutscher Wetterdienst. (2023). Wetter und Klima - Deutscher Wetterdienst - Presse - Deutschlandwetter im Jahr 2021.

Enander, A. (2021, April). Psychology of crisis and trauma. https://doi.org/10.1093/acrefore/9780190228637.013.1642

Fathi, R., & Fiedrich, F. (2022). Social Media Analytics by Virtual Operations Support Teams in disaster management: Situational awareness and actionable information for decision-makers. *Frontiers in Earth Science*, *10*, 941803. https://doi.org/10.3389/feart.2022.941803

Frewer, L. J., Miles, S., Brennan, M., Kuznesof, S., Ness, M., & Ritson, C. (2002). Public preferences for informed choice under conditions of risk uncertainty. *Public Understanding of Science*, *11*(4), 363–372. https://doi.org/10.1088/0963-6625/11/4/304

Galea, S., Nandi, A., & Vlahov, D. (2005). The epidemiology of post-traumatic stress disorder after disasters. *Epidemiologic reviews*, *27*, 78–91. https://doi.org/10.1093/epirev/mxi003

Glade, T., & Alexander, D. E. (2013). Classification of natural disasters. In *Encyclopedia of earth sciences series* (pp. 78–82). Springer Netherlands. https://doi.org/10.1007/978-1-4020-4399-4_61

Glaser, C. (2019). Selektive Wahrnehmung. In *Risiko im management* (pp. 65–68). Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-25835-1{\textunderscore}17

Goldmann, E., & Galea, S. (2014). Mental health consequences of disasters. *Annual review of public health*, *35*, 169–183. https://doi.org/10.1146/annurev-publhealth-032013-182435

Grawe, K. (1998). *Psychologische Therapie*. Hogrefe Verl. fu¨r Psychologie.

Grieve, R., Indian, M., Witteveen, K., Tolan, G. A., & Marrington, J. (2013). Face-to-face or Facebook: Can social connectedness be derived online? *Computers in Human Behavior*, *29*(3), 604–609. https://doi.org/10.1016/j.chb.2012.11.017

Gross, J. (2012). *La¨rm-induzierte Erkrankungen des Menschen*.

Jha, M. K. (2010). Natural and Anthropogenic Disasters: Vulnerability, Preparedness and Mitigation. In M. K. Jha (Ed.). Springer.

Jungha¨nel, T., Bissolli, P., Daßler, J., Fleckenstein, R., Imbery, F., Janssen, W., Kaspar, F., Lengfeld, K., Leppelt, T., Rauthe, M., Rauthe-Schöch, A., Rocek, E., M.Walawender, & Weigl, E. (2021). *Hydro-klimatologische Einordnung der Stark- und Dauerniederschla¨ ge in Teilen Deutschlands im Zusammenhang mit dem Tiefdruckgebiet Bernd vom 12. bis 19. Juli 2021*.

Kernis, M. H. (2003). TARGET ARTICLE: Toward a Conceptualization of Optimal Self-Esteem. *Psychological Inquiry*, *14*(1), 1–26. https://doi.org/10.1207/S15327965PLI1401{\textunderscore}01

Kessler, R. C., McLaughlin, K. A., Koenen, K. C., Petukhova, M., & Hill, E. D. (2012). The importance of secondary trauma exposure for post-disaster mental disorder. *Epidemiology and Psychiatric Sciences*, *21*(1), 35–45. https://doi.org/10.1017/s2045796011000758

Kreutzer, L., & Ko¨llner, V. (2020). Posttraumatische Belastungssto¨rung in der digitalen Arbeitswelt: Ein Fallbeispiel. *Die Rehabilitation*, *59*(05), 298–302. https://doi.org/10.1055/a-1148-4166

Lang, B., Dolan, R., Kemper, J., & Northey, G. (2020). Prosumers in times of crisis: Definition, archetypes and implications. *Journal of Service Management*, *32*(2), 176–189. https://doi.org/10.1108/josm-05-2020-0155

Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal, and Coping*. Springer Publishing Company.

Leon, G. R. (2004). Overview of the psychosocial impact of disasters. *Prehospital and disaster medicine*, *19*(1), 4–9. https://doi.org/10.1017/s1049023x00001424.

Liu, B., & Austin, L. (2010). *How publics use social media to communicate during crises: Proposing the social-mediated crisis communication model*.

Lo¨hr, K., Weinhardt, M., & Sieber, S. (2020). The "World Cafe´" as a Participatory Method for Collecting Qualitative Data. *International Journal of Qualitative Methods*, *19*, 160940692091697. https://doi.org/10.1177/1609406920916976

Makwana, N. (2019). Disaster and its impact on mental health: A narrative review. *Journal of family medicine and primary care*, *8*(10), 3090–3095. https://doi.org/10.4103/jfmpc.jfmpc{\textunderscore}893{\textunderscore}19

Mason, V., Andrews, H., & Upton, D. (2010). The psychological impact of exposure to floods. *Psychology, Health & Medicine*, *15*(1), 61–73. https://doi.org/10.1080/13548500903483478

Mayring, P. (1991). Qualitative Inhaltsanalyse. In U. Flick (Ed.), *Handbuch qualitative sozialforschung* (pp. 209–213). Psychologie-Verl.-Union.

Mu¨ller, F., Tomczyk, S., Fathi, R., von Berg, M.-L., Tutt, L., & Fiedrich, F. (2023). Soziale Medien als psychosoziale Ressource in Krisen und Katastrophen. *Mensch und Computer 2023 – Workshopband*. https://doi.org/10.18420/MUC2023-MCI-WS01-348

Müller-Cyran, A., & Zehentner, P. (2013). *Wenn der Tod plo¨tzlich kommt: Vom Umgang mit Schicksalsschla¨gen - Das Kriseninterventionsteam im Einsatz*. Heyne.

Nestler, S. (2017). Fla¨chendeckende Kommunikation im Stromausfall durch regionale IKT.

Nikendei, A. (2012). *Psychosoziale Notfallversorgung (PSNV): Praxisbuch Krisenintervention*. Stumpf + Kossendey.

Norris, F. H., Stevens, S. P., Pfefferbaum, B., Wyche, K. F., & Pfefferbaum, R. L. (2008). Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *American journal of community psychology*, *41*(1-2), 127–150. https://doi.org/10.1007/s10464-007-9156-6

Paton, D., & Flin, R. (1999). Disaster stress: An emergency management perspective. *Disaster Prevention and Management: An International Journal*, *8*(4), 261–267. https://doi.org/10.1108/09653569910283897

Paulsen, N., Weiß, R., & Wo¨stmann, L. (2023). *Mehr als 50 Millionen Deutsche nutzen soziale Medien* (Bitkom e. V., Ed.). https://www.bitkom.org/Presse/Presseinformation/Mehr-als-50-Millionen-Deutsche-nutzen-soziale-Medien

Pausch, M. J., & Matten, S. J. (2018). Symptome der PTBS. In *Trauma und Traumafolgesto¨rung* (pp. 27–52). Springer, Wiesbaden. https://doi.org/10.1007/978-3-658-17886-4{\textunderscore}4

Pearlin, L. I. (1989). The Sociological Study of Stress. *Journal of Health and Social Behavior*, *30*(3), 241. https://doi.org/10.2307/2136956

Plapp, S. T. (2003). *Wahrnehmung von Risiken aus Naturkatastrophen: Eine empirische Untersuchung in sechs gefa¨hrdeten Gebieten Su¨d- und Westdeutschlands* [Dissertation]. Universita¨t Fridericiana zu Karlsruhe.

Quarantelli, E. (Ed.). (2005). *What is a Disaster?* Routledge. https://doi.org/10.4324/9780203984833

Radio Euskirchen. (2022). Als die Flutwelle Nordrhein-Westfalen im Juli 2021 traf.

Rao, K. (2006). Psychosocial support in disaster-affected communities. *International Review of Psychiatry*, *18*(6), 501–505. https://doi.org/10.1080/09540260601038472

Reuter, C., & Kaufhold, M.-A. (2017). Fifteen years of social media in emergencies: A retrospective review and future directions for crisis informatics. *Journal of Contingencies and Crisis Management*, *26*(1), 41–57. https://doi.org/10.1111/1468-5973.12196

Rufat, S., Tate, E., Burton, C. G., & Maroof, A. S. (2015). Social vulnerability to floods: Review of case studies and implications for measurement. *International Journal of Disaster Risk Reduction*, *14*, 470–486. https://doi.org/10.1016/j.ijdrr.2015.09.013

Schreier, M. (2014). Varianten qualitativer Inhaltsanalyse: Ein Wegweiser im Dickicht der Begrifflichkeiten. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *Vol 15*, No 1 (2014). https://doi.org/10.17169/FQS-15.1.2043

Schro¨der, L. (2022). Die Flutkatastrophe 2021: Ein Protokoll (WDR, Ed.).

Simon, T., Goldberg, A., & Adini, B. (2015). Socializing in emergencies—a review of the use of social media in emergency situations. *International Journal of Information Management*, *35*(5), 609–619. https://doi.org/10.1016/j.ijinfomgt.2015.07.001

Solomon, S. D., & Green, B. L. (1992). Mental health effects of natural and human-made disasters. *PTSD RESEARCH QUARTERLY*, (Volume 3, Number 1), 1–8.

Steier, F., Gyllenpalm, B., Brown, J., & Bredemeier, S. (2008). Politische Beteiligung. In N. Kersting (Ed.). VS Verlag fu¨r Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91071-0 10

Steigleder, S. (2008). *Die strukturierende qualitative Inhaltsanalyse im Praxistest: Eine konstruktiv kritische Studie zur Auswertungsmethodik von Philipp Mayring*. Tectum-Verl.

Tajfel, H., & Turner, J. C. (2004). The Social Identity Theory of Intergroup Behavior. In J. T. Jost & J. Sidanius (Eds.), *Political psychology* (pp. 276–293). Psychology Press, an imprint of Taylor and Francis. https://doi.org/10.4324/9780203505984-16

Tempest, E. L., Carter, B., Beck, C. R., & Rubin, G. J. (2017). Secondary stressors are associated with probable psychological morbidity after flooding: A cross-sectional analysis. *European Journal of Public Health*, *27*(6), 1042–1047. https://doi.org/10.1093/eurpub/ckx182

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science (New York, N.Y.)*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Umeda, M., Chiba, R., Sasaki, M., Agustini, E. N., & Mashino, S. (2020). A Literature Review on Psychosocial Support for Disaster Responders: Qualitative Synthesis with Recommended Actions for Protecting and Promoting the Mental Health of Responders. *International Journal of Environmental Research and Public Health*, *17*(6), 2011. https://doi.org/10.3390/ijerph17062011

UNDRR & CRED (Eds.). (2020). Human cost of disasters: An overview of the last 20 years (2000-2019).

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Vitaliano, P. P., Maiuro, R. D., Bolton, P. A., & Armsden, G. C. (1987). A psychoepidemiologic approach to the study of disaster. *Journal of Community Psychology*, *15*(2), 99–122. https://doi.org/10.1002/1520-6629(198704)15:2{\textless}99::AID-JCOP2290150202{\textgreater}3.0.CO;2-Q

Walker-Springett, K., Butler, C., & Adger, W. N. (2017). Wellbeing in the aftermath of floods. *Health & place*, *43*, 66–74. https://doi.org/10.1016/j.healthplace.2016.11.005

Zisgen, J., Kern, J., Thorm, D., & Ertl, T. (2014). *#Hochwasser - Visuelle Analyse von Social Media im Bevo¨lkerungsschutz*.



Visions for Future Crisis Management

Track Chairs

Anne-Marie Barthe-Delandë, IMT Mines Albi

Aurélie Montarnal, IMT Mines Albi

Rob Grace, Texas Tech University

Technology Enhanced Actual Situation Awareness Model for Dog-Handlers Working with Search and Rescue Dogs

Stella Polikarpus
Estonian Academy of Security Sciences
stella.polikarpus@sisekaitse.ee

Kristine Steen-Tveit
Centre for Integrated Emergency Management (CIEM), University of Agder
kristine.steen-tveit@uia.no

Ana María Aldea Reyes
Spanish School of Rescue and Detection with Dogs (ESDP)
esdp.eu@escuelasalvamento.org

Marta Burgos González
ESDP
m.burgos@escuelasalvamento.org

Susana Izquierdo Funcia
Spanish School of Rescue and Detection with Dogs (ESDP)
s.izquierdo@escuelasalvamento.org

Graham Edgar
University of Gloucestershire
gedgar@glos.ac.uk

ABSTRACT

Rescue dogs are used in disaster response because of their ability to find victims using their excellent sense of smell, supported by dynamic decision-making skills. Disaster area environment is hazardous for both dogs and dog-handlers. Therefore, new technology is being developed to assist. In this article we argue that if new technologies are developed with the aim of making search and rescue missions in disaster zones safer and more successful, then there is a need to research how such technology can support communication between dog, handler, and incident command; in particular, how is shared situation awareness developed and supported. We present a model that scopes the interactions between human, dog, and technology together with advanced techniques for assessing those interactions through measurement of (shared) situation awareness and understanding. With such a wide range of potential technologies (such as VR, AR, 360-degree online video streaming, gas detecting sensors and special communication vests for dogs) having a clear view of how different parts of a team interact, and how to assess those interactions, is crucial. The use of technology (e.g. AR) is not limited to use in the implementation of search and rescue procedures, but can also be used in testing, and training for, other new technologies.

Keywords

Search and Rescue Dogs (SAR Dogs), Dog-Handlers, Quantitative Analysis of Situation Awareness (QASA), Technology Enhanced Actual Situation Awareness Model for Dog-Handlers working with Search and Rescue Dogs (TEASAM-DH & SAR Dogs)

INTRODUCTION

Every year man-made or natural disasters occur all-over the world. New technologies are constantly developed to improve the safety and effectiveness of first responders, while also saving the lives of victims. New technologies including virtual (and/or augmented) reality, robots and advanced sensors are being tested to replace dogs and humans operating in hostile environments (Sanfilippo & Rañó, 2023). Because dogs have long been established as excellent partners for first responders with the ability to operate autonomously while looking for victims in

disaster areas, robots are usually built to mimic dogs' capabilities in the disaster area, such as the ability to move over rough terrain (Bellicoso et al., 2018). At the same time technology enhanced vests are designed to be used with rescue dogs and their handlers (Arnold et al., 2019; Kasnesis et al., 2022). New wearable devices are being developed that allow sensing, geolocation, and monitoring of the state-of-health of the search and rescue (SAR) dog. Such devices also support dog-handler communication, including sending commands to the dog (Contini & Martins, 2024). While such new technology has the potential to be of great benefit, it is important to conduct research to establish whether such benefits include, for example, improved situation awareness (SA) for first responders and incident commanders (Roldán et al., 2017).

Dogs, with roughly 40 times more smell-sensitive receptors than humans (Kumar, 2022), rely heavily on their sense of smell as their primary sense. Currently, no robots or sensors can match a dog's sense of smell. Unlike the human brain, which prioritizes a large visual cortex, the dog's brain is primarily dominated by an olfactory cortex (Correa, 2011). A dog has more than 100 million olfactory receptors in its nose (see Table 1), while humans have only 5 million (Padodara & Jacob, 2014). Dogs (and humans) also have a vomeronasal organ (also called Jacobson's organ) which is located above the palate and behind the upper incisors (Correa, 2011), that is particularly sensitive to pheromones. Dogs' sense of smell has led to them being trained for search and rescue missions (SAR missions), to search for victims in disasters or missing people around the world – and to detect for, amongst other things, narcotics, drugs and contraband agricultural products, homicide victims and forensic materials from corpses.

Table 1. Scent-Detecting Cells in Humans and Dog Breeds
(adopted from Padodara & Jacob, 2014 who reference Stanley and Sarah, 2013)

Species	Number of Scent Receptors
Humans	5 million
Dachshund	125 million
Fox Terrier	147 million
Beagle	225 million
German Shephard	225 million
Bloodhound	300 million

Urban search and rescue (USAR) teams are governed internationally by regulations published by INSARAG (International Search and Rescue Advisory Group). It is mandatory that all rescue teams include technical search units for the search and detection of victims under collapsed structures in earthquakes and other disasters. Currently, there is no more effective and efficient mechanism for searching for victims than trained dogs. These dogs are the main tool of the technical search teams of USAR groups. Given the current effectiveness of dogs within USAR teams, it is essential that any new technologies should be integrated appropriately, and the success of such integration assessed. We propose a model to support the development and integration of the technologies in such a way that the co-operation of dogs and their handlers improves and that, crucially, such improvement is manifest in (and can be assessed by) improved situation awareness.

We use a definition of situation awareness (SA) first proposed by Endsley (1995), that SA is the: "perception of the elements in the environment within a volume of time and space, comprehension of their meaning, and the projection of their status in the near future". In the event of a disaster, SAR dogs and their handlers (as well other first responders, inhabitants, and victims) need to quickly be aware of elements of the disaster such as the location (GPS coordinates) of SMS-warnings, sirens, drones, collapsed buildings, safe areas etc. Individuals and teams within the disaster zone then need to be able to integrate the available information to comprehend the situation and, at highest level of SA, to predict the most effective and safe activity in that situation. While technology has the potential to enhance SA at all levels and for all involved, it is crucial not to assume that the mere adoption of new technological tools will automatically bolster SA; thorough research is necessary. Additionally, it is important to emphasize that the implementation of any new technology requires the adjustment of standard operational procedures, and the training of both dog-handlers and dogs, to effectively integrate those new technologies.

MEASURING SITUATION AWARNESS

While the general concept of SA has become well established and researched in a number of domains, the use of simulation-based training scenarios as a method to improve SA has been the subject of relatively little research (Polikarpus et al., 2019, 2022). Nevertheless, Model-based Design is suggested for complex systems when real-

time simulation are needed (Bélanger et al., 2010) and a Venn diagram is the world's most recognizable graphic approach to visualize and learn from similarity and matches (Koll et al., 2023). We combined our practical experiences of training SAR dogs and dog handlers using technology to measure SA into Technology Enhanced Actual Situation Awareness Model for Dog-Handlers presented as a Vienn-diagram below in the paper.

Endsley (Endsley, 2021) reviewed a number of methods for measuring SA one of which, Quantitative Analysis of Situation Awareness (QASA – although the attribution in Endsley's paper is incorrect) has the unique feature of being able to measure *negative* SA – indicating not only that somebody has no awareness of the situation, but that they also misunderstand it. In disaster situations, under heavy stress, such misunderstanding could easily happen to the extent that the dog-handler does not understand that his/her dog is threatened by the hazard from environment.

QASA measures both actual situation awareness (ASA) and actual situation understanding (ASU). See Table 2 for an overview of abbreviations and descriptions. ASA and ASU are rather novel terms used in recent projects, such as the FireFront project (Thoelen et al., 2020), compared to Endsley`s well cited SA definition. ASA refers to an awareness of the 'ground truth'. That is, the situation as it objectively really is. Good SA is, arguably, a necessary, but not sufficient, criterion for effective performance. An individual (or team) needs to not only be aware of the situation, but they also need to understand it - ASU.

QASA also measures *perceived* SA (PSA) and perceived SU (PSU). PSA/U refers to how good an individual *believes* their awareness/understanding of the situation to be. It is not guaranteed that ASA/U and PSA/U will necessarily match, and the cause of some catastrophic accidents can be traced back to a mismatch of ASA/U and PSA/U. It is therefore desirable to assess not only ASA and ASU, but also PSA and PSU. Good ASA can tell you that the technology is doing its job in delivering the necessary information, ASU can tell you whether the information is presented in such a way that the end user can understand it. PSA/U can tell you whether the individual has an *awareness* of their own performance level. Given the range and sophistication of the technique, we therefore propose the use of QASA (Edgar et al., 2018) to measure aspects of SA and SU while using The Collaborative Authoring Process Model for Virtual Simulations (CAPM) (Polikarpus et al., 2022) to assess the impact of new technologies on SA and SU, as shared across teams.

Table 2. Overview of abbreviations and descriptions

Abbreviations	Description	Reference
SA	Situation Awareness	Endsely, 1995
SU	Situation Understanding	Thoelen et al., 2020
PSA	Perceived Situation Awareness	Edgar et al., 2018
ASA	Actual Situation Awareness	Thoelen et al., 2020
ASU	Actual Situation Understanding	Thoelen et al., 2020
PSU	Perceived Situation Understanding	Thoelen et al., 2020
QASA	Quantitative Analysis of Situation Awareness	Edgar et al., 2018

MEASURING DOG-HANDLER COMMUNICATION AND DOG SITUATION AWARENESS

Dogs have been a part of human society for longer than any other domestic species (Benz-Schwarzburg et al., 2020), and they have a certain level of SA/U, although it may not be as sophisticated as that of humans. As discussed, SA/U involves understanding and interpreting what is happening in the environment, and dogs exhibit this ability in various ways. Research has shown that dogs are capable of assessing their surroundings, using their senses (including sight, hearing, and smell) to recognize familiar people and objects, and responding to changes in their environment (Benz-Schwarzburg et al., 2020).

The real bond between humans and dogs is marked by cohabitation in the same living space (Kuhne, 2016). Research has suggested that canine domestication has uniquely endowed dogs with two essential abilities for cooperative dog-human problem-solving: social tolerance and social attentiveness. These capabilities enable dogs to adapt their behavior to that of their human partners (Ostojić & Clayton, 2014). Collaboration necessitates such partners dedicating ample attention to one another to adapt or coordinate their actions. Additionally, social learning involves focusing on a demonstrator's actions and the context in which that action occurs (Huber et al., 2009).

In the case of human-dog interactions, the level of attentiveness is significantly influenced by the relationship between the dog and the human (Horn et al., 2013). Dogs have proven successful in several tasks that are thought to require high attention towards humans, as demonstrated by experiments on social learning, social reflecting, communication, responding to unequal rewards, and cooperation (Benz-Schwarzburg et al., 2020). Based on these processes, dogs make informed decisions about proper behavior (Huber, 2016). Studies have investigated the

effects of training methods on the performance of working dogs e.g.(Alexander et al., 2011) while others have emphasized the role of the dog-handler relationship e.g. (Lefebvre et al., 2007). In their study, Lefebvre et al. (2007) noted that working dogs residing in their handler's household, and participating in sporting activities together, exhibited increased sociability, obedience, and efficiency. The bond between humans and canines was found to have a positive impact on the dog's focus on its handler, leading to improved communication and, in turn, heightened performance among working dogs. Figure 1. shows a SAR dog, alert and focused on its handler, patiently awaiting instructions. Further, Szetei et al., (2003) found that in their study that, when locating a hidden object, SAR dogs may favor human communicative signals, such as pointing, over olfactory information.



Figure 1. SAR dog waiting for instructions (Private photo Kristine Steen-Tveit)

While searching for possible victims the tasks of the dog could be described in the following steps:

1. **Searching** for scents in the area of interest;
2. **Detecting** the location of the first relevant odor particles in an area;
3. **Locating** the origin of the desired odor, or the place with the highest concentration of odor particles;
4. **Manifestating** the appropriate behaviour - the dog decides the area on which it will focus its attention and then manifests the appropriate behavior to demand the handler's attention.

The first three steps could be seen as closely related to situation awareness levels: the dog collects the (sensory) information from the area, comprehends the meaning (picking out the appropriate odour to indicate a person in distress), and then makes the correct response in summoning the handler.

In step four, fixation or manifestation, barking is the most used alerting behavior for dogs in disaster areas. Currently, the barking alert is essential in the process of searching for victims in collapsed structures, as there is no technological tool that is sufficiently precise and safe to replace the barking alert. When the dog finds the person, s/he must indicate the position of the victim to the handler either by remaining in the immediate vicinity of the person and barking until the handler arrives – or use his/her teeth to acquire the "bringsel" hanging from his/her collar. The dog then returns to the handler carrying the “bringsel” in the mouth (see Figure 2) (Hiljanen & Nordin, 2012) and, on receiving a specific command, will show the handler the location of the victim.



Figure 2. Dog with bringsel (Private photo: Kristine Steen-Tveit)

With the addition of GPS technology, however, the dog could “mark the victim” by sitting or lying down and, if the GPS indicates that the dog stays in position for more than 5 seconds, a signal would be sent to the dog-handler.

Another way in which technology could be utilized would be to develop technology that detects the dog’s activity and barking in real-time, and alerts the dog handler that the dog has found a victim, and where the dog (and victim) is located (Kasnesis et al., 2022).

Given that there is an appropriate behaviour for the dog at any particular point in the search process, we could measure dog and dog-handler ASA and ASU by measuring the dogs’ actual behavior against the optimum behavior (considered further below). More conventional measures could be used to measure handler ASA and ASU.

The dog handler stops a dog working via voice or whistle (vile) command, but this means that means that the dog must stay within earshot of the handler. Some GPS collars now have the capability to emit a sound or vibration initiated from the handler's GPS device, and which the dogs can be trained to interpret as a "find owner" signal. Currently, however, all the devices developed to locate dogs during a search are not viable to locate dogs during their work under rubble, since the devices lack reliability and safety. Research shows that if dog is more than 60 meters from handler the contact breaks (Contini & Martins, 2024). Therefore, still now if the dog locates a victim out of sight of the handler (under debris or in a hidden place), the only way to signal the discovery to the handler is by means of an audible signal (barking) or by means of a dog body signal (remaining in position such as sitting or lying or taking bringsel into mouth). If these dogs’ signals could be interpreted by a technical tool that then sends a warning signal to the dog-handler, it has a potential to speed up the search and rescue process and improve dog-handler ASA. For safety purposes, such technology need not be mounted on the dog. The dogs’ behaviours could be monitored remotely by networked drones or robots, potentially vastly increasing the ‘area coverage’ for

any one handler and dog. In terms of SA, the 'perception' aspect of the handler is greatly increased.

If technology is to be successfully used with SAR dogs, then another crucial area of research is to study how dogs interact with that technology. The authors have noticed, for example, that dogs, without any training to interact with technology, will ignore everything on a television show - except animal shows. Such shows appear to draw the dog's attention and lead to the dog approaching the TV (see Figure 3). There is research that explores the ability of canines to interact with touchscreen interfaces in order to improve service dogs' communication with their handlers (Zeagler, Zuerndorfer, et al., 2016). Animal-computer interaction research has enhanced the creation of computer-aided systems that allow handlers and dogs to work together better while conducting SAR tasks (Zeagler, Byrne, et al., 2016).



Figure 3. Dog watching animal shows on TV (Private photo: Stella Polikarpus)

Extending the idea (discussed above) of using technology to simply detect a dog barking, there is research on using artificial intelligent (AI) to convert dog barks into English (Adamson, 2018) or to allow the dog to send signals using a worn vest in emergency situations (Kasnesis et al., 2022; Zeagler, Byrne, et al., 2016). Dogs have been trained to alert another human to a problem using a vest that includes a handy mechanical lever which, when pulled by the dog, plays an audio message such as: "My handler needs you to come with me!" (National Geographic Kids). AI can also be used to assess the behavior of dogs in a particular situation (Farhat et al., 2023) which, when compared with identified optimal behavior would provide a semi-automated, and real time, measure of ASA and ASU.

MEASURING THE ACTUAL SA OF HUMANS AND DOGS WITH THE HELP OF TECHNOLOGY

Technology mediated systems lend themselves well to knowledge-based measures of SA as there is often a clear, and recorded, path of information flow around members of the team (including IT systems). One such knowledge-based technique, as discussed above, is QASA (Edgar et al., 2018) that has been successfully used to assess ASA, PSA, ASU, and PSU in simulated safety critical situations (See table 2 for a reminder of descriptions). QASA is based on probing team members SA/SU (and this could include IT systems) using true/false (T/F) statements drawn from the situation, combined with an underlying analysis based around signal detection theory (SDT). While this approach works well for human and (potentially) IT elements of a team, the training overhead in getting dogs to respond to T/F probes is likely to be extreme. There are, however, performance-based measures of SA e.g. (Burge & Chaparro, 2012) or (Collier & Follesoe, 1995) that are also based around signal detection theory and so would dovetail well with QASA. These measures compare actual performance with optimum performance and could (in theory!) be readily adapted to assessing the performance of dogs (and hence their underlying SA/SU) in a safety critical context such as SAR.

If methods of assessing SA/U in a team can be developed then those methods could be used to assess the effect of introducing technologies such as smart phones, GPS positioning, 360º camera streaming, advanced sensors, robots or drones, dog vests, artificial intelligent, machine learning, augmented reality, local data networks, radios, and much more. As the interactions of dogs and handlers are increasingly mediated by technology, the effectiveness of that technology can be assessed by measuring SA/U. Different technologies may impact different aspects of SA. For example, improved communication systems (between, dog, handler, and incident command) (Arnold et al., 2019) may improve SA. The use of VR and AR (for example by handler) for data fusion and presentation could potentially increase SU. Assessing such effects is vital. One thing that additional technology almost always brings is additional information (e.g. 360º camera streaming is proposed instead of images processing (Arnold et al., 2019)). This could be an excellent support for SA (and SU) or it could overload the individual causing a loss of SA/U (a handler too busy trying to understand the camera output and missing a signal from the dog, for example). More is not always better.

The use of dogs in disaster areas raises traditional questions of ethics, such as whether dogs should be used in hazardous areas involving gas leaks, bombed buildings etc. Introducing technology, however, raises a whole new range of issues around such things as data protection, data transfer, and other cyber security issues. It is therefore necessary to scope the SAR task in terms of the elements (human, dog, technology) and the interactions between them. It would be helpful to have a simple model for crisis management research that embodies the areas of

overlap between elements of a system and the different facets of SA (and SU) that are generated by those overlaps. Such a model allows the assessment of new technologies to be targeted and interpreted as a part of the 'big picture. In Figure 4. we propose a Technology Enhanced Actual Situation Awareness Model for Dog-Handlers working with SAR Dogs. For simplicity, at this stage, we have only scoped for SA (not SU, although the regions would be similar).

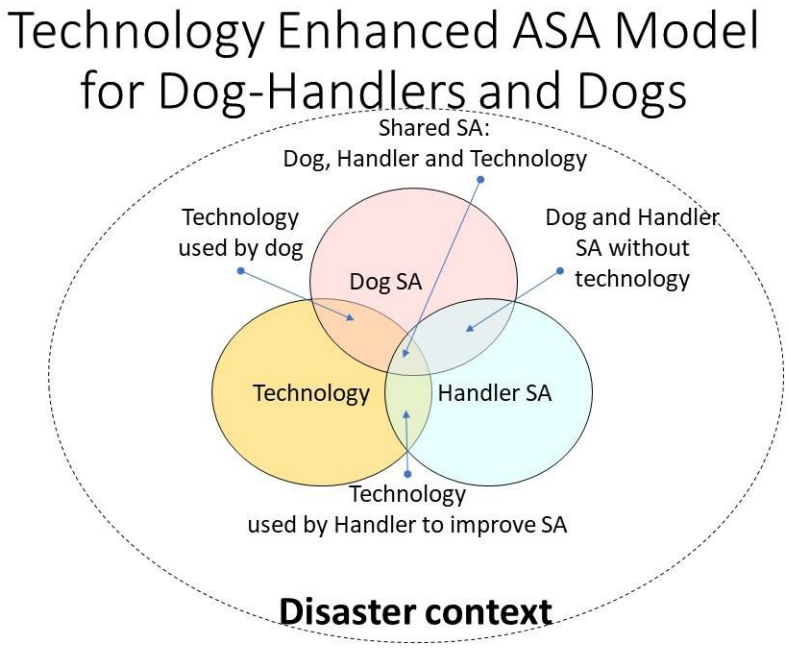


Figure 4. Technology Enhanced Actual Situation Awareness Model for Dog-Handlers working with Search and Rescue Dogs

In the model (Figure 4.) seven different areas can be identified; three discrete areas and four overlapping. The first discrete area is the dog with its unique set of characteristics, experiences and training that help it to make dynamic decisions in specific context. The second is the dog-handler who has same factors influencing his/her decision-making as dog but with additional factors and possibilities to gain situation awareness from command center. The third is the technology like 360-degree video streaming, GPS, communication, and observation technology as well as AI. Each element will develop its own awareness (SA) of the situation. For example, technology can have false positive alert in image recognition (Arnold et al., 2019). There are then three dual overlaps. The first dual overlap is the dog and handler, here belongs behaviours dog and handler use to communicate with each other, and standard operation procedures based on what dogs are certified. The second the dog and technology where most used have been vests with different technologies attached to it (GPS positioning, sensors, communication tools) but also dogs might learn to use specific technology themselves to communicate like touchscreens (Zeagler, Zuerndorfer, et al., 2016). The third component is the handler and technology, and dog handlers should be able to use all technological tools and information systems needed to gain "big picture" and follow standard operational procedures.

These overlaps will mediate the SA of each of the overlapping elements and, one would hope, improve SA rather than impair it. For example, introducing technological communication systems might improve the handler SA (they could receive information from other handlers, the 'big picture' from incident command, and so on). The last area in Figure 4. is a triple overlap between handler, dog, and technology and this is, effectively, where the 'shared SA' of the team (handler, dog, technology) resides. It is, perhaps, in this area where the disaster context specific technologies could have most impact on enhancing the effectiveness and safety of both dogs and handlers.

FROM CURRENT PRACTICES TO FUTURE STRATEGIES: ADVANCING CANINE TRAINING

Regardless of the type of SAR activities, canine training is extensive. Training programs are conducted in many countries worldwide, such as Germany, Japan, Australia, and Norway. Although the training processes in different countries often share similarities, different countries and regions may have specific protocols and standard operational procedures. Standards for the certification of SAR dogs also vary by certifying organization. The

National Association for search and rescue (NASAR) has, however, established voluntary qualification standards for, among other roles, disaster search (Jones et al., 2004).

The overarching goal of all training is to prepare the dogs and their handlers for effective SAR operations in different disaster contexts. The canine training process is a rigorous process that requires dedication, time, and specialized skills and includes obedience training, scent discrimination, search techniques, trailing, and avalanche/disaster training. The foundation of SAR dog training is socialization. Dogs are frequently brought out into public settings to get as accustomed to a wide range of situations and people as possible. Since leash control is crucial in SAR, obedience, and the capacity to obey directed instructions are also crucial. For dogs that are trained to search in both urban and wilderness settings, athleticism and agility are essential. Dogs are taught to move boldly, but cautiously, over surfaces that they would normally avoid. Urban SAR canines train extensively on debris, which might include broken concrete, old cars, uneven terrain, chain link fences, and rebar. Urban search and rescue dogs are required to climb ladders, negotiate tunnels, and rappel with their handler (see photos in Figure 5.). As opposed to patrol dogs, aggression towards humans or other animals is not tolerated in SAR dogs (Jones et al., 2004).

The future of SAR canine training is likely to see a significant integration of technology to enhance the safety of both the dogs and their handlers, as well as to improve SA. Examples can be smart vests for dogs equipped with sensors that can provide real-time data on the dog's position, vital signs, temperature, and overall well-being (Contini & Martins, 2024; Kasnesis et al., 2022). Exploring ways to enhance communication between SAR dogs and their handlers through technology could improve the efficiency and effectiveness of search operations. Integrating technologies into canine training requires a balance between innovation and reliability and a means of assessing whether inclusion of that technology has been beneficial. It has been already observed in the research that new training requirements must be developed to integrate effectively technologies to be used by service dogs (Zeagler, Zuerndorfer, et al., 2016). Additionally, considerations for ethical and humane treatment of the dogs, as well as the need for adaptability in various operational contexts, will be essential in shaping the future of SAR dog training.



Figure 5. Urban SAR dog training. (Private photo: Marita Poulsen)

CONCLUSION

It is important that the context and characteristics of dogs and their handlers are taken into account when new technologies are introduced with the aim of improving ASA and, by extension, dogs and handlers' safety, communication, performance, standard operating procedures, ease of standardization etc. Improved performance of the team means being better able to assist victims in different disaster areas. The proposed Technology Enhanced Actual Situation Awareness Model for Dog-Handlers working with SAR Dogs is provided as a way of scoping out the interactions and issues inherent in any complex team; and a team involving humans, dogs, and new technology, is likely to be complex. The approach in this paper proposes the model as a way of scoping those complex interactions, combined with advanced techniques for measuring SA/U to assess the impact (both good and bad) of such interactions. Such assessment will necessarily involve possibly disparate communities such as technology developers, researchers, and practitioners. The overall aim is, however, to further improve the effectiveness of SAR dogs through the use of technology, and the ISCRAM community has all the competencies needed to achieve this.

REFERENCES

Adamson, A. (2018). *Pet Translator: Scientist Developing Device To Convert Dog Barks Into English Language*. Tech Times. https://www.techtimes.com/articles/218841/20180115/pet-translator-scientist-developing-device-to-convert-dog-barks-into-english-language.htm

Alexander, M. Ben, Friend, T., & Haug, L. (2011). Obedience training effects on search dog performance. *Applied Animal Behaviour Science*, *132*(3–4), 152–159. https://doi.org/10.1016/j.applanim.2011.04.008

Arnold, S., Ohno, K., Hamada, R., & Yamazaki, K. (2019). An image recognition system aimed at search activities using cyber search and rescue dogs. *Journal of Field Robotics*, *36*(4), 677–695. https://doi.org/10.1002/rob.21848

Bélanger, J., Venne, P., & Paquin, J.-N. (2010). The What, Where and Why of Real-Time Simulation. *Planet Rt 1.1*, 37–49.

Bellicoso, C. D., Bjelonic, M., Wellhausen, L., Holtmann, K., Günther, F., Tranzatto, M., Fankhauser, P., & Hutter, M. (2018). Advances in real-world applications for legged robots. *Journal of Field Robotics*, *35*(8), 1311–1326. https://doi.org/10.1002/rob.21839

Benz-Schwarzburg, J., Monsó, S., & Huber, L. (2020). How Dogs Perceive Humans and How Humans Should Treat Their Pet Dogs: Linking Cognition With Ethics. *Frontiers in Psychology*, *11*(December). https://doi.org/10.3389/fpsyg.2020.584037

Burge, R., & Chaparro, A. (2012). The Effects of Texting and Driving on Hazard Perception. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 715–719. https://doi.org/10.1177/1071181312561149

Collier, S. G., & Follesoe, K. (1995). SACRI: A Measure of Situation Awareness for Nuclear Power Plant Control Rooms. *Experimental Analysis and Measurement of Situation Awareness*, 115–122.

Contini, M. S., & Martins, L. E. G. (2024). Development of a wearable device to provide electronic assistance to search and rescue dogs. *Research on Biomedical Engineering*, *40*(1), 265–280. https://doi.org/10.1007/s42600-024-00341-y

Correa, J. E. (2011). *The Dog 's Sense of Smell*. https://ourdogssavelives.org/wp-content/uploads/2012/02/Dogs-Sense-of-Smell.pdf

Edgar, G., Catherwood, D., Baker, S., Sallis, G., Bertels, M., Edgar, H. E., Nikolla, D., Buckle, S., Goodwin, C., & Whelan, A. (2018). Quantitative Analysis of Situation Awareness (QASA): modelling and measuring situation awareness using signal detection theory. *Ergonomics*, *61*(6), 762–777. https://doi.org/10.1080/00140139.2017.1420238

Endsley, M. R. (1995). Measurement of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 65–84. https://doi.org/10.1518/001872095779049499

Endsley, M. R. (2021). A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM. *Human Factors*, *63*(1), 124–150. https://doi.org/10.1177/0018720819875376/FORMAT/EPUB

Farhat, N., Lazebnik, T., Monteny, J., Moons, C. P. H., Wydooghe, E., van der Linden, D., & Zamansky, A. (2023). Digitally-enhanced dog behavioral testing. *Scientific Reports*, *13*(21252), 1–12. https://doi.org/10.1038/s41598-023-48423-8

Hiljanen, T., & Nordin, S. (2012). *FCI guidlines for International Nordic Style Trials* (pp. 1–22). https://www.fci.be/medias/UTI-REG-SNO-en-503.pdf

Horn, L., Huber, L., & Range, F. (2013). The Importance of the Secure Base Effect for Domestic Dogs – Evidence from a Manipulative Problem-Solving Task. *PLoS ONE*, *8*(5), e65296. https://doi.org/10.1371/journal.pone.0065296

Huber, L. (2016). How Dogs Perceive and Understand Us. *Current Directions in Psychological Science*, *25*(5), 339–344. https://doi.org/10.1177/0963721416656329

Huber, L., Range, F., Voelkl, B., Szucsich, A., Virányi, Z., & Miklosi, A. (2009). The evolution of imitation: what do the capacities of non-human animals tell us about the mechanisms of imitation? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2299–2309. https://doi.org/10.1098/rstb.2009.0060

Jones, K. E., Dashfield, K., Downend, A. B., & Otto, C. M. (2004). Search-and-rescue dogs: an overview for veterinarians. *Journal of the American Veterinary Medical Association*, *225*(6), 854–860.

https://doi.org/10.2460/javma.2004.225.854

Kasnesis, P., Doulgerakis, V., Uzunidis, D., Kogias, D. G., Funcia, S. I., González, M. B., Giannousis, C., & Patrikakis, C. Z. (2022). Deep Learning Empowered Wearable-Based Behavior Recognition for Search and Rescue Dogs. *Sensors*, *22*(3), 993. https://doi.org/10.3390/s22030993

Koll, O., von Wallpach, S., & Uzelac, B. (2023). Stakeholder-oriented brand management: A Venn-diagram approach to monitor brand associations. *European Management Journal*, *41*(3), 437–444. https://doi.org/10.1016/j.emj.2022.03.004

Kuhne, F. (2016). Behavioural responses of dogs to dog-human social conflict situations. *Applied Animal Behaviour Science*, *182*, 38–43. https://doi.org/10.1016/j.applanim.2016.05.005

Lefebvre, D., Diederich, C., Delcourt, M., & Giffroy, J.-M. (2007). The quality of the relation between handler and military dogs influences efficiency and welfare of dogs. *Applied Animal Behaviour Science*, *104*(1–2), 49–60. https://doi.org/10.1016/j.applanim.2006.05.004

National Geographic Kids. (n.d.). *Scientists Invent Device that Lets Dogs Talk (Yes, Really !)*. National Geographic Kids. Retrieved February 13, 2024, from https://www.natgeokids.com/uk/discover/animals/general-animals/scientists-device-dogs-talk/

Ostojić, L., & Clayton, N. S. (2014). Behavioural coordination of dogs in a cooperative problem-solving task with a conspecific and a human partner. *Animal Cognition*, *17*(2), 445–459. https://doi.org/10.1007/s10071-013-0676-1

Padodara, R. J., & Jacob, N. (2014). Olfactory Sense in Different Animals. *The Indian Journal of Veterinary Medicine*, *2*(1), 1–14. https://www.researchgate.net/publication/262932824

Polikarpus, S., Bøhm, M., & Ley, T. (2019). Training Incident Commander's Situational Awareness—A Discussion of How Simulation Software Facilitate Learning. In T. Väljataga & M. Laanpere (Eds.), *Digital Turn in Schools—Research, Policy, Practice* (pp. 219–234). Springer, Singapore. https://doi.org/10.1007/978-981-13-7361-9_15

Polikarpus, S., Ley, T., Hazebroek, H., Edgar, G., Sallis, G., Baker, S., & Masip, A. F. (2022). Authoring virtual simulations to measure situation awareness and understanding. In H. Karray, A. De Nicola, N. Matta, & H. Purohit (Eds.), *ISCRAM 2022 Conference Proceedings – 19th International Conference on Information Systems for Crisis Response and Management* (pp. 428–433). https://idl.iscram.org/files/stellapolikarpus/2022/2430_StellaPolikarpus_etal2022.pdf

Roldán, J., Peña-Tapia, E., Martín-Barrio, A., Olivares-Méndez, M., Del Cerro, J., & Barrientos, A. (2017). Multi-Robot Interfaces and Operator Situational Awareness: Study of the Impact of Immersion and Prediction. *Sensors*, *17*(8), 1720. https://doi.org/10.3390/s17081720

Sanfilippo, F., & Rañó, I. (2023). Mimicking the Sense of Smell of Search and Rescue (SAR) Dogs: a Bio-inspired Steering Framework for Quadruped Robots. *Proceedings of the International ISCRAM Conference*, *2023-text*(May), 892–901. https://doi.org/10.59297/fbhp3870

Szetei, V., Miklósi, Á., Topál, J., & Csányi, V. (2003). When dogs seem to lose their nose: an investigation on the use of visual and olfactory cues in communicative context between dog and owner. *Applied Animal Behaviour Science*, *83*(2), 141–152. https://doi.org/10.1016/S0168-1591(03)00114-X

Thoelen, F., Vastmans, J., Blom Andersen, N., Boehm, M., Holm, L., Arendtsen, B., Polikarpus, S., Taukar, M., Kütt, T., Fikke, R. C., Geertsema, T., Hazebroek, J. C., Tonnaer, C., Weewer, R., Figueras Masip, A., Fuste Castella, R., Catherwood, D., Baker, S., Brookes, D., … Walker, S. (2020). FireFront: A new tool to support training in Fireground Situation Awareness, Situation Understanding and Bias. *International Fire Professional*, *34*, 34–39.

Zeagler, C., Byrne, C., Valentin, G., Freil, L., Kidder, E., Crouch, J., Starner, T., & Jackson, M. M. (2016). Search and rescue: Dog and handler collaboration through wearable and mobile interfaces. *ACM International Conference Proceeding Series*, *15-17-Nove*. https://doi.org/10.1145/2995257.2995390

Zeagler, C., Zuerndorfer, J., Lau, A., Freil, L., Gilliland, S., Starner, T., & Jackson, M. M. (2016). Canine computer interaction: Towards Designing a Touchscreen Interface for Working Dogs. *Proceedings of the Third International Conference on Animal-Computer Interaction*, *15-17-Nove*, 1–5. https://doi.org/10.1145/2995257.2995384

Exploring the roles and challenges of disruptive technologies for emergency management in Aotearoa New Zealand

Lukas Kroher
Otto-Friedrich University Bamberg
lukas.kroher@gmail.com

Marion Lara Tan
Massey University
M.L.Tan@massey.ac.nz

Raj Prasanna
Massey University
R.Prasanna@massey.ac.nz

ABSTRACT

This work-in-progress paper presents preliminary findings from an exploratory workshop on disruptive technologies in emergency management in Aotearoa New Zealand. Thematic analysis identified three main themes on the role of disruptive technologies: (1) creating a common operating picture, (2) enhancing interaction, and (3) organising information. Despite the transformative potential of disruptive technologies, there are several challenges, including reliability and consistency of information, managing large volumes of data, integration and interoperability, and security and privacy concerns over data usage, among others. The study highlights the need for robust collaboration and resource-sharing mechanisms to leverage disruptive technologies in New Zealand's emergency management landscape. Future work will investigate these identified themes through additional interviews and workshops involving emergency management professionals. Further focus will examine the preliminary findings to the broader concept of situational awareness in emergency contexts.

Keywords

Disruptive technologies, emergency management, thematic analysis, exploratory study

INTRODUCTION

Aotearoa New Zealand (NZ) faces a range of natural hazards, including geophysical and hydrometeorological hazards. Earthquakes, a significant hazard, register approximately 20,000 events annually, with around 33 exceeding a magnitude of 5.0, posing risks to infrastructure (GeoNet, 2023). In the hydrometeorological context, flooding is the most frequently occurring hazard, projected to occur every eight months in NZ due to climate change implications (Mason et al., 2021). Cascading impacts could also trigger further hazards such as tsunami and landslides.

The consequences of hazards extend beyond the physical environment and can also have social impacts. As Goffman (1974, p. 379) aptly describes in his framework analysis, such negative experiences can result in a phenomenon he refers to as "reality anomically flutters". This phenomenon extends beyond individuals, manifesting on a macro level in the face of natural hazards, triggering a process of renegotiation and normalisation during highly destructive events. Such processes are pivotal, especially when emergency management teams and decision-makers synchronise their various perceptions with on-site realities. Therefore, efficient handling of information plays a key role in creating a picture of the situation that is as reliable as possible (Jayawardene et al., 2021). This process can be facilitated by technologies that sit in a moderating position between the individual and the situation (Hörl, 2011).

Previous studies have shown that technology can be important in the complex and dynamic emergency management system to support effective and timely decisions (Endsley, 1995; Prasanna et al., 2017). However, technology keeps evolving. Emerging and disruptive technologies such as artificial intelligence (AI), the Internet of Things (IoT), and Big Data are changing many fields, including disaster risk reduction and management (ITU,

2019). This evolution enables faster processing and sharing of critical information, adapting to the changing nature of disaster risks (Munawar et al., 2022). In this rapidly changing landscape, it is imperative that we update and advance research on exploring how technologies can be leveraged to enhance disaster risk reduction.

This study investigates disruptive technologies in emergency management in NZ. This work-in-progress paper presents preliminary findings from a workshop held in Auckland, NZ, on 7 November 2022, in alignment with the ISCRAM Asia Pacific Conference. The objectives of this study are twofold: (1) to investigate the role of technologies in the production process of generating the necessary information for responding to a disaster, and (2) to examine the challenges in using the information for the subsequent response process. Note that the data gathering for this study was conducted before major flooding events in January and February 2023. [1] Therefore, it is not within the scope of this study to include lessons learned from these events.

BACKGROUND

This section delves into key concepts underpinning the study and provides a brief overview of the current research landscape, crucial for understanding the interplay between disruptive technologies and emergency response.

Key Concepts

These terms are defined to provide context and assist consistency in addressing the topic of this study of disruptive technologies in emergency management.

- **Disasters** are multifaceted events causing damage, losses, and disruption to societies or communities (Peek et al., 2021). As defined by the United Nations Office of Disaster Risk Reduction, a disaster is a "A serious disruption of the functioning of a community or a society at any scale due to hazardous events interacting with conditions of exposure, vulnerability and capacity, leading to one or more of the following: human, material, economic and environmental losses and impacts." (UNDRR, 2017)
- **Disruptive technologies** are new technologies that can give rise to a new avenue of doing business or research (Munawar et al., 2022). In this study's context, these are technological developments that can provide potential solutions for improvements in emergency management (ITU, 2019).
- **Emergency management** encompasses the organisation and management of resources for addressing all aspects of emergencies, including preparedness, response, and recovery (UNISDR, 2009).

Current Research on the Topic

In recent years, emergency management has witnessed significant disruptions facilitated by technological advancements (Munawar et al., 2022). Academic publications in the past five years (2019 onwards) have frequently highlighted the pivotal role of various technologies in reshaping the landscape . Noteworthy technologies frequently mentioned in recent literature include Artificial Intelligence (AI), big data, sensors, Internet-of-Things (IoT), autonomous vehicles including unmanned aerial vehicles (UAV), blockchain, and fog/edge computing (see Table 1). It's important to note while Table 1 summarises the applications for these disruptive technologies and outlines the associated challenges, the list is not exhaustive. These examples provide a snapshot of prevalent technologies discussed in the field.

From the literature, these technologies can play a crucial role across the entire disaster lifecycle. However, the integration of disruptive technologies also has challenges. According to Munawar et al. (2022), one significant challenge is the absence of policies, especially on data governance. Much of the needed information (e.g. real-time data, infrastructure details) lies within city or municipal governance and policies, which may be hard to access. Data integration protocol is also needed to ensure the new technologies are aligned to provide information from proper data gathering and modelling (Akter & Wamba, 2019). For example, for a disaster management system to use big data, an enormous amount of location-specific data may come in during a disaster; the system should be capable of aiding agencies to usher this big data to become useful information to provide real-time actionable services to the people in need (Akter & Wamba, 2019).

[1] The January Auckland flooding and Cyclone Gabrielle were significant natural hazard events for NZ in 2023. insurance costs for these events are estimated at NZ$1.65 billion (ICNZ, n.d.) These events also resulted in loss of human life, with four people killed in the Auckland floods (RNZ, 2023a) and eleven people losing their lives during Cyclone Gabrielle (RNZ, 2023). Many individuals were left stranded and hundreds needed to be rescued from rooftops.

Table 1. Summary of frequently discussed disruptive technologies, applications, and challenges

Technology	Examples of applications	Challenges	References
Artificial intelligence	• Accelerate disaster management across all phases, e.g., geospatial data analysis for disasters. • Hazard assessment (e.g. cyclone intensity, floods, and landslides) • Impact assessment (e.g. infrastructure damage) • Crowd management and situation analysis, prediction, and evacuation planning • Sentiment analysis: find opinions or emotions on a specific topic	• Requires precise data with effective feature selection. • Large amounts of data and accuracy of data. • Fragmented efforts on narrow use cases without broader stakeholder collaboration; AI tools underutilized in decision-making. • Data sources may be inaccessible or incompatible with other datasets; data may not captured systematically.	(Arora et al., 2023; Munawar et al., 2022)
Big data (including social media) and cloud services	• Use of GIS and GPS for integrated hazard analysis with large datasets. • Use of social media and location data to find trends and threats for real-time monitoring (e.g. bushfires) • Helps identify critical events and needs assessment • Gauge public opinion to influence decision-making and offer early feedback on government policies.	• Protecting data confidentiality, privacy, and ownership • Storage and processing systems may struggle with performance delivery • Capability, including infrastructure & personnel • Availability, scalability, quality & interoperability.	(Akter & Wamba, 2019; Arora et al., 2023; Bania et al., 2023; Munawar et al., 2022)
Autonomous vehicles (e.g. UAV)	• Images captured by drones can be used for impact assessment to evaluate the situations.	• Battery, physical load, processing power, difficult manoeuvring conditions	(Munawar et al., 2022)
Blockchain	• Blockchain in disaster management aids collaboration through peer-to-peer communication and smart contracts. • Blockchain's sharing economy boosts resource management accountability and auditability. • System democratisation enables multi-disciplinary involvement whilst increasing the quality of decision-making.	• Lack of trust or understanding amongst stakeholders on blockchain technology.	(Cheklin et al., 2022; Krichen et al., 2024; Vermiglio et al., 2022)
Sensors, internet-of-things (IoT)	• IoT paired with AI enables early prediction using real-time data. • IoT tracks equipment in disaster management. • Sensors and IoT aid data collection, analytics, and early warnings. • Wireless autonomous networks assist communication without pre-existing infrastructure. • Supports both centralized and decentralized network architectures.	• The heterogeneous nature of interconnected devices poses several interoperability challenges.	(Arora et al., 2023; Bania et al., 2023; Munawar et al., 2022; Prasanna et al., 2022)
Fog/edge computing	• Ability to process data in real-time, closer to end-user devices, resulting in reduced response times without the intervention of the cloud; enables quick access to data and faster decision-making.	• Limited resources at the edge. • Data communication challenges.	(Bania et al., 2023; Militano et al., 2023; Prasanna et al., 2022)

With disruptive technologies, significant concerns such as security and privacy arise, as data availability also increases the risk of manipulation and unauthorised use (ITU, 2019). Another significant challenge lies in the high implementation and maintenance costs of these cutting-edge technologies (Akter & Wamba, 2019; Cheklin et al., 2022). There is a need for cost-effective solutions that deliver long-term benefits.

Much of the existing literature focuses on testing and providing sample use cases, but further research is needed to understand their implementation and performance during and post-disaster (Munawar et al., 2022). Understanding the role of collaboration is a research opportunity when utilising technologies. Closer collaboration facilitating effective sharing of resources is needed to utilise disruptive technologies for emergency management (Cheklin et al., 2022). Disruptive technologies cannot be fully utilised or scaled up when there is incomplete data sharing and ineffective use of resources (Cheklin et al., 2022).

Acknowledging the existing work, gaps, and opportunities in this area, this paper investigates more closely how emergency managers and technology researchers view the needs and challenges of utilising disruptive technologies for emergency management in NZ.

METHOD

Data Acquisition

A workshop was conducted to understand the technology needs and challenges for emergency management in NZ. The workshop was held on 7 November 2022 in alignment with the second ISCRAM Asia Pacific Regional Conference with the theme "Dealing with the Unexpected". The conference was held at RMIT University in Melbourne, Australia, with a satellite venue at AUT in Auckland, NZ. The workshop was held in Auckland to reach the target NZ-based participants.

The workshop was designed with two objectives in mind: (1) to investigate the role of technologies in the production process of generating the necessary information for responding to a disaster, and (2) to examine the challenges in using the information for the subsequent response process. With these, the workshop focused on exploring the technological desires and requirements of the emergency management sector. Particularly through a large-scale crisis scenario centred on situation awareness and decision-making needs. The workshop was held with subject matter experts (academics) and emergency management practitioners. The participants were formed into four focus groups to investigate the patterns of *doing-crisis* approaches. This method helps to connect unrelated individuals (and thus perspectives) from various backgrounds. Furthermore, the fluid nature of focus groups might trigger knowledge, experiences and emotions more effectively than face -to-face interviews (Witkowski et al., 2021). The workshop obtained peer-reviewed approval (Project number: 4000026842) under the ethical conduct code of Massey University for research with human participants deemed "low risk".

Participant Recruitment and Composition

Invitations for the workshop were dispatched to the Civil Defence Emergency Management (CDEM) groups, the National Emergency Management Agency (NEMA), scholars in disaster research and ISCRAM Conference attendees. This procedure ensured that academics (13 participants) and practitioners (18) were adequately represented in this focus group. Specifically, the latter group contains members from regional councils (8), representatives from different ministries (3), employees from NEMA (4), and participants from the lifelines sector (1) and the private technology sector (1). There were no admission prerequisites except for being physically present in Auckland. Conducting an in-person workshop in Auckland presents a limitation, as it could potentially introduce bias through the barrier of limiting attendance for those living further away.

Workshop Structure and Guide Questions

The workshop is centred around a hypothetical emergency scenario, which was presented to the participants: The Greater Wellington region is affected by long and strong ground shaking due to an earthquake (Magnitude 8.9, with shaking up to 10 MMI in Wellington) on the south-eastern coast of NZ's North Island. In addition to the damage caused by the tremors, the earthquake unleashed a series of tsunami waves that caused widespread flooding throughout the affected area.

The participants were then formed into groups, with each group consisting of academics and practitioners. Each group was tasked with the (imaginary) coordination of a national intelligence task force in their position as emergency management leaders during the post-disaster response phase in the first week. Against this, the focus group deliberated on tackling this situation for two hours, based on five guide questions aligning with the research objectives (see Table 2). The questions revolve around informa tion and technology—questions 1 and 2 focus on

efficiently handling information in response to the event. Questions 3, 4, and 5 revolve around using technology as part of the response strategies to the hypothetical earthquake.

Table 2. Guide questions to stimulate the discussion

Topic	Question	Aligning to objectives
Focus on information	1. What are the *information/intelligence needs* to build situation awareness for the initial first week to achieve your goals for response and recovery	(1) to investigate the role of technologies in the production process of generating the necessary information for responding to a disaster
	2. What *challenges* do you expect to arise when acquiring and using information during the first week of response and recovery?	(2) to examine the challenges in using the information for the subsequent response process
Focus on technology	3. What technological tools and solutions are *currently* available (in practice or research) to address the information needs and problems?	(1) and (2)
	4. What *new* technological tools and solutions can be developed to enhance situation awareness?	(2)
	5. What technologies should be *prioritised* to be researched, re-evaluated, and developed?	

Data Analysis

Qualitative research methodology is employed to analyse the workshop data. Following this epistemological understanding, the experiences of each speaker contribute to institutionalising a shared, intersubjective reality (concept originating from Berger et al., 1967). Braun and Clarke (2006) suggest a reflexive thematic analysis (TA) to convey the general trends of focus group discussions. As per their suggestion, single themes could be understood as "pattern(s) of shared meaning" (Braun et al., 2019, p. 845) that are arranged around a central concept. In consequence, the (1) themes are detached from single speaker positions, and the (2) themes emerge through (a) discourses within the focus groups and (b) through the active, moulding role of the researcher (Braun et al., 2019). Responses were not coded or differentiated according to the speaker's identity or affiliation due to the methodological preconditions of the thematic analysis and privacy concerns. The thematic analysis follows a six-step process, as detailed in Table 3.

Table 3. Thematic analysis (based on Braun et al., 2019)

Phase	Explanation/ justification	Application
Transcription	Make the data material accessible for systematic analyses	AI-supported conversion from audio to text
Familiarisation	Getting used to data	Rereading, making casual notes
Coding	A systematic process of making sense of the data	Deductive coding of the semantic meaning of each segment (through the guide questions)
Constructing themes	Merge codes with similar meanings to provisional themes	Unstructured, inductive organisation of codes around similar semantic meanings (independent of pre-given structure)
Revising themes	Define and establish themes within itself and in their relation to each other.	Rigorous merging of code groups; investigating similarities on latent meaning within each group, leading to a differentiation process to other themes.
Producing report	Final coherence check how well themes can be augmented	Presenting the final themes to a scientific audience

The thematic analysis started with transcribing the recording of each focus group discussion with the assistance of AI technology. Following a brief familiarisation period with the database, the coding process began by tackling the transcribed data sequentially through each guide question, and every segment was allocated a label denoting its semantic meaning. The deductive themes guided by the workshop questions are illustrated in Figure 1.[2] The

[2] This includes the initial segments from Q1 to Q4. As Q5 does not contain any new aspects (it only highlights

codes were then categorised based on comparable explicit sense, where analysing the data was not limited to the guide question structure in this process. No longer limited to pre-defined structures, the codes were aggregated and re-organised by collapsing and combining the codes with similar semantic meanings.

Further into the thematic analysis process, the research team identified three main themes (See Figure 2), which were then reviewed to ensure clear definitions and distinctions. This also necessitated the generalisation of meaning by transforming the themes into a more implicit and latent level. The methodological approach used in this study is thus aligned with the principles of grounded theory, which includes theoretical sampling, transcription, coding, and contrasting (Bryant & Charmaz, 2007)[3].

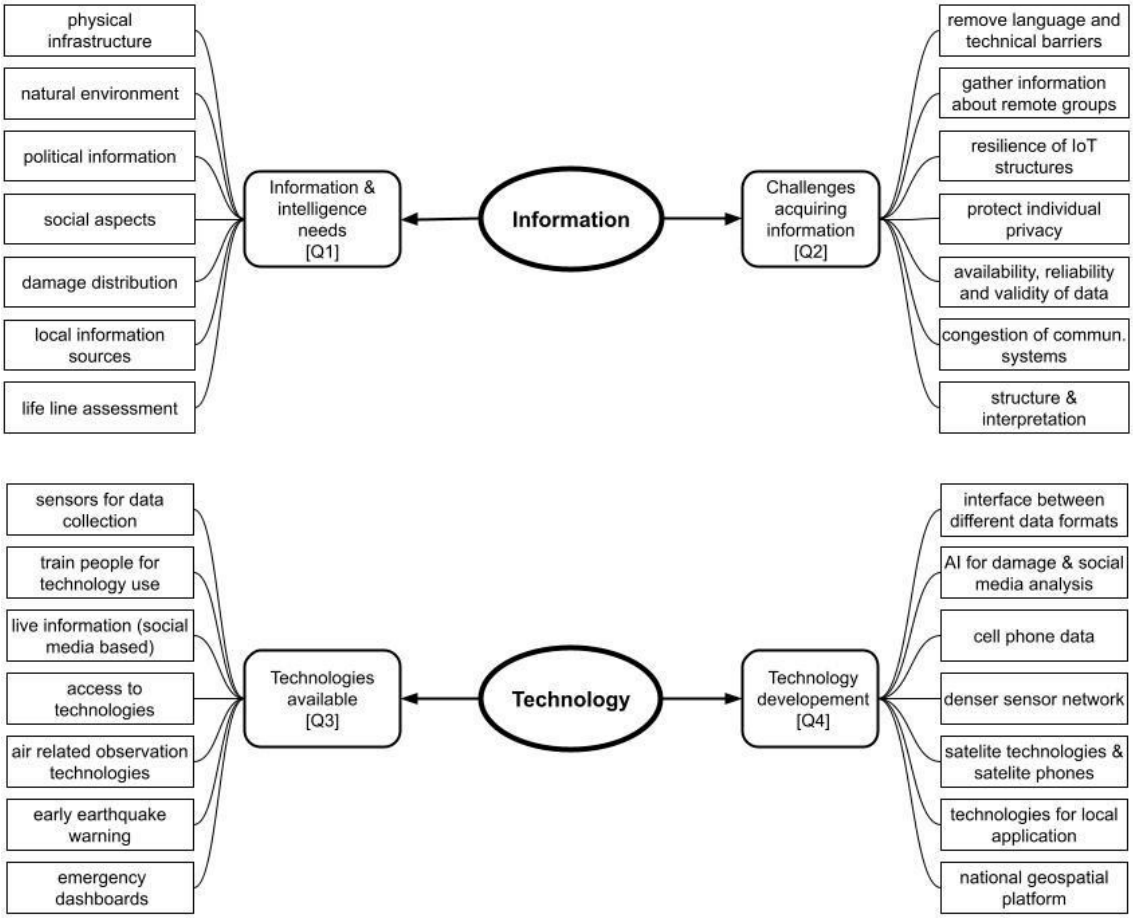


Figure 1. Initial themes deducted from guide questions

PRELIMINARY RESULTS

Through the production process of generating necessary information for response, the preliminary findings show that there are three main themes for disruptive technologies' role in emergency management: (1) the common operating picture (COP), (2) the interaction (IA), and finally, (3) the organisation of information (OI). Figure 2 illustrates the graphical composition of the resulting thematic analysis. Alongside these themes, we highlight some of the challenges encountered.

Common Operating Picture (COP)

The first theme is *associated with achieving a stable information status* for all involved in the emergency

aspects already mentioned as important), it is not included.

[3] Guest et al. (2012, p. 13) stress, that the Thematic analysis "method does not preclude theoretical development", which is also part of the grounded theory idea. However, he does not clarify whether this implies a totally autonomous macro theory or a middle-range proposition. The latter could be provided by the thematic analysis during the generalisation step.

management processes; this concept was commonly referred to by the participants as the 'Common Operating Picture'. The participants desire a reliable and statically rendered common picture to make sound decisions. Statically rendered means information that is consistent and with little ambiguity. The following participant's quote highlights the detailed questions that can arise from responders to help develop COP for decision-making:

> *Are the roads open or shut? My top priority is to know whether the state highways are open and then the local roads. Next, are the ports impacted? Can [they] roll on, roll off? […] Can gear ships and other kinds of ships operate? What's the impact on the fuel system? From the wharves at the ports to the tank farms, which are normally low-lying at [parts]? […] Impacts on service stations and […] the impacts on the airports, […] all of these tell us what the impacts on the supply chain are. Can food be delivered, and water and response activities be done?* (Group C participant)

When discussing the COP theme, the participants were particularly interested in learning about the status of lifelines, the physical environment, and the social aspects of the situation. Their interest in lifeline assessment looks into information about the availability of essential goods; this includes power, fresh water, food and medical support in hospitals. The physical aspects of COP relate to knowing more about the built environment, including evaluating critical infrastructure, such as accessibility to ports, roads, railways, and airports. The COP theme also involves learning about the disaster situation's social aspects, covering the damage distribution from a social needs perspective.

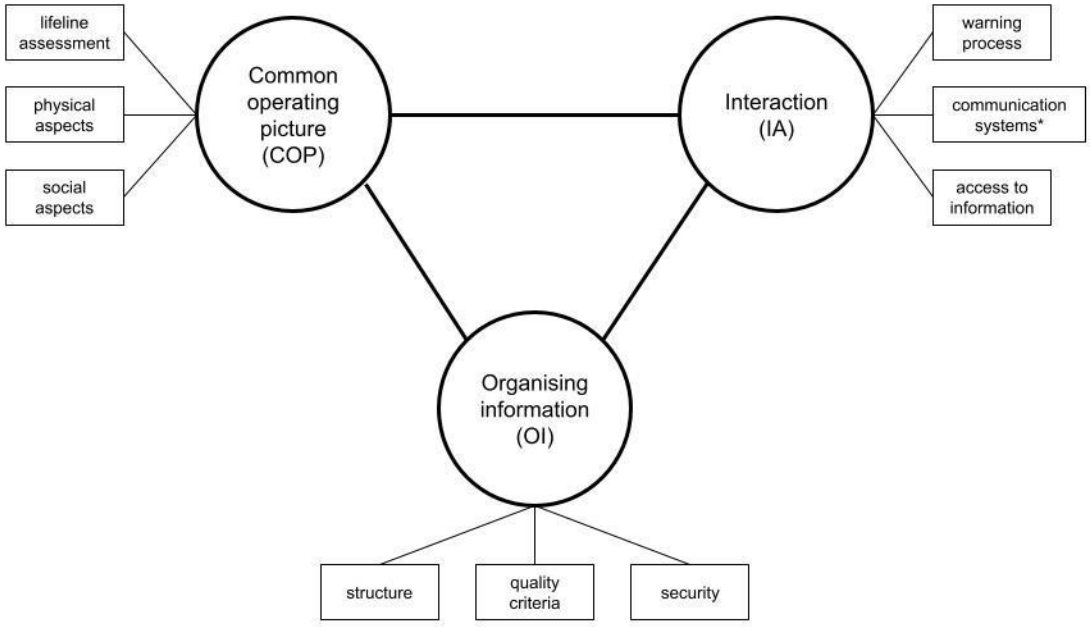


Figure 2. Visualisation of the preliminary results from the thematic analysis

To build a COP, participants acknowledge the importance of detailed information, which may be voluminous and come from multiple sources. Participants recognise the dynamic and emerging nature of disasters and thus multifaceted. Therefore, disruptive technologies' role is to gather reliable data and ensure the stability of the picture to withstand and integrate new players and emerging information. Hence, COP, although desired to be stable, must be adaptable, always to construct a more reliable awareness of the current situation. The technologies participants identified relating to COP include remote sensing and resilient sensor networks to aid impact assessments. The participants also mentioned the need to have the ability to process large amounts of information in quicker timeframes (i.e. AI and big data).

Interaction (IA)

The IA theme describes a *dynamic information exchange process* within and between the emergency managers and affected people. This flow of information can involve different technologies and revolves around multiple stakeholders, including providing information within and from the authorities and ultimately interactin g with the

information from those affected. Quotes from the participants highlight the interaction theme with emergency managers and the affected people:

> *I've just said one of the initial challenges is going to be actually around communications and actually being able to share data and share intelligence. Yeah, because it's going to be significant losses of infrastructure like fibre breakages, yeah* (Group C Participant).
>
> *I guess how accessible is it both from […] a physical device but [also] is it in a way [in a] language you understand or a format you understand?* (Group B Participant)

When the discussion relates to the IA theme, it often revolved around interactions regarding the overall warning process, the role of communication systems, and the accessibility of information. An effective *warning process* is essential to give people the best chance to take protective actions. Participants highlighted that the process is not just technological but involves human and organisational processes. A working *communication system* is the core basis for all interaction during emergencies. Therefore, its maintenance is a crucial aspect of emergency response and thus holds great priority among workshop participants. Participants emphasised that excessive communication volume can pose a considerable challenge to the system, especially on the initial day following the disaster. This escalates the likelihood of system failure due to congestion during this critical phase. Participants highlighted that disruptive technologies can be the source of congestion (e.g., multiple da ta points, including social media) and a solution (e.g., IT systems incorporating AI, big data, and cloud services).

Participants highlighted multiple examples of disruptive technologies that aid in providing access to crucial information. This includes satellite imagery, use of UAVs, sensor networks, and social media. The ability to receive and communicate this message also relies upon access to suitable technical apparatus. On a societal lev el, access to information lies in reducing barriers to interpreting warnings and instructions, for example, eliminating language barriers and designing warnings with minimal complexity for ease of comprehension. The distribution and receipt of information also depends on various facts, among others, such as age (i.e., digital divide) and way of life (i.e., urban vs. rural). Access to crucial information serves as the basis for engaging with affected individuals.

Organisation of Information (OI)

The OI theme describes the *all-encompassing mechanism to store, analyse, and extract justifiable meaning* from incoming data. This includes data gathered automatically or through human input. Particular topics highlighted in the OI theme include: structure of incoming data, aspects and challenges around quality standards, and a secure handling of sensible information. Quote from participants highlights the OI theme:

> *I think definitely AI is coming. […] so this can help us to quickly identify whether, for example, the building has been damaged or it is not […] So you have texts, pictures or also other 3D geospatial data, so you need different type of technologies to process them. But I think, in general, AI I think can help us to reduce the human efforts.* (Group B Participant)

Different data types become available during emergencies and interpreting them will require effort. For example, for social media, data types include pictures, text, or geotagged information. Even geographical data can be presented in different formats (e.g. polygons vs. points vs. lines). Participants highlighted that in NZ, there is no national consensus on how to manage data, so every institution (e.g. different regional councils) maintains their own data collection. As a result, participants highlighted the role of disruptive technologies in standardising data types and making them interoperable. It was mentioned that AI holds a significant role in structuring incoming data. Therefore, AI-based methods such as targeted filtering for specific keywords and hashtags (for social media analysis), natural language processing, and sentiment analysis can be used. [4]

The quality of the information is another significant topic for OI. Quality includes the data's reliability (trustworthiness) and validity (no bias). Evaluating data quality can strain human resources during a crisis; this is where AI can help cross-check various data formats and sources to enhance the quality of interpretations. The evaluation will also ascertain which councils or platforms offer the most accurate information on the physical environment and infrastructure as part of the quality review process.

[4] Note that the workshop was held before the public release of ChatGPT on 30 November 2022 (Lock, 2022).

Finally, since the data can come from multiple sources and may contain data about human beings, attention is also needed to the security aspects. Participants also highlighted the importance of honouring data privacy concerns.

DISCUSSION

In answering the paper's first objective of investigating the role of technologies in generating information for response, the results have shown multiple insights into disruptive technologies' role in emergency management through the three themes: COP, OI, and IA. These themes, in practice, are not isolated but influence each other, shaping decision-making processes during emergencies: forming a reliable COP relies on the structured interpretation of data within the OI, which informs the coordination of interactions (IA) during emergencies. COP is a critical bridge between OI and IA, facilitating informed and strategic decision-making in crisis scenarios.

To fulfil the second objective of the paper on examining the challenges in using the information for the subsequent response process, we have found several challenges relating to COP, OI, and IA. There are challenges related to the reliability and stability of the COP due to the dynamic nature of disasters and the need for consistent yet adaptable information. Managing large volumes of diverse data from multiple sources and integrating emerging information poses significant difficulties. Interaction (IA) challenges include maintaining effective communication and data sharing among stakeholders, managing technology congestion, and ensuring information accessibility and interpretation for diverse populations. Organisation of Information (OI) challenges encompass managing data structure and quality, validating data authenticity, and addressing security and privacy concerns. Aside from the themes, general challenges involve achieving interoperability across different platforms, mitigating human resource strain through AI-assisted data evaluation, and navigating ethical and security concerns associated with data usage and integration.

Aligning with the challenges highlighted in the workshop, recent events such as the Auckland Floods and Cyclone Gabrielle that occurred in NZ (after the workshop was conducted) have highlighted the vulnerability of existing, more traditional technologies to provide sustained sensing and communication networks during and post-disaster scenarios (Bush International Consulting, 2024). These incidents have exposed the limitations of current systems, particularly in maintaining continuous data transmission and providing good situational awareness for decision-makers. The disruption to sensing networks and communication channels during such events poses significant challenges for frontline responders and wider emergency management authorities in making timely and informed decisions (Delgado-Ferro et al., 2022).

Compared to more traditional technology-based systems, disruptive technologies have significant potential to aid disaster management, as highlighted in the literature (e.g. Munawar et al., 2022) and workshop discussions. While existing technology-based systems support various functional roles of emergency management and the associated tasks, they still face several limitations that may hinder their efficiency and effectiveness. Operationalising traditional technology-driven systems is often constrained by high costs, limited scalability, centralised architectures, and vulnerable technology infrastructure, hindering timely and effective disaster response and recovery efforts (Prasanna et al., 2022). Therefore, ensuring the reliability of our technological systems and infrastructure is paramount and foundational. Regarding post-event infrastructure management, connectivity, serviceability, and restoration are key elements to consider (Kongar et al., 2017).

Robust communication systems and infrastructure are fundamental components of effective emergency management. This includes the need for an adaptive but reliable COP, a working system for IA that incorporates the whole warning process to ensure communities have access to information, and for OI to have a structure that ensures data quality and security. However, despite its capabilities, the successful integration of disruptive technologies hinges on robust policy frameworks that facilitate stakeholder collaboration and resource sharing. However, realising the full potential of disruptive technologies in NZ necessitates addressing challenges related to incomplete data sharing and ineffective resource utilisation.

Limitations

Study limitations warrant consideration. Method limitations, such as the potential biases due to the in-person format and the workshop location in Auckland, may have influenced the workshop outcomes. Future research should address these limitations by expanding the scope of data collection, including holding multiple more data gathering options, including workshops and interviews with online options to attend, to ensure comprehensive coverage of relevant stakeholders and perspectives.

Temporal constraints imposed by the workshop's timing may limit the comprehensiveness and applicability of the workshop findings due to technological developments and disaster events. Notably, the workshop was held before the public release of ChatGPT, a significant development in generative AI, which could potentially alter perspectives on AI's utility in emergency scenarios. Furthermore, data gathering for this study was conducted

before major flooding events in NZ during January and February 2023. Subsequent data collection efforts with subject matter experts and emergency manager practitioners could provide valuable insights into the changes and adaptations following these events.

Future Work

While this study provides preliminary insights into the COP, OI, and IA themes and the associated challenges, future investigation by the research team will delve deeper into the sub-themes in this study and explore the intricate interconnections between COP, OI, and IA in greater detail. Comparative analysis between COP and the broader concept of situational awareness could provide valuable insights into their overlapping and distinct aspects, further enriching our understanding of information management in emergency contexts. Future work by the research team includes conducting follow-up interviews or workshops with more emergency managers in NZ.

Moreover, future studies should focus on use cases in simulated or testing environments. Implementation and evaluation studies would provide insightful, actionable recommendations for enhancing NZ's emergency management practice and policy development. Furthermore, social science studies could explore opportunities to study collaboration and partnerships between researchers, practitioners, and policymakers. By translating the insights gained from this research into practical interventions and policy recommendations, stakeholders can effectively address challenges related to data sharing, resource utilisation, and technology integration in emergency management.

REFERENCES

Akter, S., & Wamba, S. F. (2019). Big data and disaster management: a systematic review and agenda for future research. *Annals of Operations Research*, *283*(1–2), 939–959. https://doi.org/10.1007/s10479-017-2584-2

Arora, S., Kumar, S., & Kumar, S. (2023). Artificial Intelligence in Disaster Management: A Survey. In *Lecture Notes in Networks and Systems* (Vol. 552). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-6634-7_56

Bania, A., Iatrellis, O., & Samaras, N. (2023). Information Communication Technologies (ICTs) and Disaster Risk Management (DRM): Systematic Literature Review. In *Lecture Notes in Intelligent Transportation and Infrastructure*. Springer Nature Switzerland. https://doi.org/10.1007/978 -3-031-23721-8

Berger, P. L., Luckmann, T., & Light, D. W. (1967). The social construction of reality : a treatise in the sociology of knowledge. *American Sociological Review*, *32*(137). https://doi.org/10.2307/3710424

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic Analysis. *Handbook of Research Methods in Health Social Sciences*, 843–860. https://doi.org/10.1007/978-981-10-5251-4_103

Bryant, A., & Charmaz, K. (2007). *The SAGE Handbook of Grounded Theory*. SAGE Publications Ltd. https://doi.org/https://doi.org/10.4135/9781848607941

Bush International Consulting. (2024). *Hawke's Bay Civil Defence and Emergency Management Group Respones to Cyclone Gabrielle*. https://www.hbemergency.govt.nz/assets/Uploads/HBCDEM-Response-to-Cyclone-Gabrielle-Final-Report.pdf

Cheklin, M., Naeni, L. M., & Killen, C. (2022). Opportunities for Application of Disruptive Technology in a Disaster Management System to Address Gaps in Australian Bushfire Response. *45th AUBEA Conference*, 282–292.

Delgado-Ferro, F., Navarro-Ortiz, J., Chinchilla-Romero, N., & Ramos-Munoz, J. J. (2022). A LoRa WAN Architecture for Communications in Areas without Coverage: Design and Pilot Trials. *Electronics (Switzerland)*, *11*(5). https://doi.org/10.3390/electronics11050804

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society* , *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

GeoNet. (2023). *Earthquake Statistics*. https://www.geonet.org.nz/earthquake/statistics_long

Goffman, E. (1974). Frame analysis: An essay on the organization of experience. In *Frame analysis: An essay on the organization of experience*. Harvard University Press.

Guest, G., MacQueen, K., & Namey, E. (2012). *Applied Thematic Analysis*. https://doi.org/10.4135/9781483384436

Hörl, E. (2011). *Die technologische Bedingung: Beiträge zur Beschreibung der technischen Welt* . Suhrkamp Verlag.

ICNZ. (n.d.). *Cost of natural disasters*. Insurance Council of New Zealand. https://www.icnz.org.nz/industry/cost-of-natural-disasters/

ITU. (2019). Disruptive technologies and their use in disaster risk reduction and management. In *International Telecommunication Union*.

Jayawardene, V., Huggins, T. J., Prasanna, R., & Fakhruddin, B. (2021). The role of data and information quality during disaster response decision-making. *Progress in Disaster Science*, 100202. https://doi.org/10.1016/j.pdisas.2021.100202

Kongar, I., Esposito, S., & Giovinazzi, S. (2017). Post-earthquake assessment and management for infrastructure systems: learning from the Canterbury (New Zealand) and L'Aquila (Italy) earthquakes. *Bulletin of Earthquake Engineering*, *15*(2), 589–620. https://doi.org/10.1007/s10518-015-9761-y

Krichen, M., Abdalzaher, M. S., Elwekeil, M., & Fouda, M. M. (2024). Managing natural disasters: An analysis of technological advancements, opportunities, and challenges. *Internet of Things and Cyber-Physical Systems*, *4*(September 2023), 99–109. https://doi.org/10.1016/j.iotcps.2023.09.002

Lock, S. (2022). *What is AI chatbot phenomenon ChatGPT and could it replace humans* . The Guardian. https://www.theguardian.com/technology/2022/dec/05/what-is-a i-cha tbot-phenomenon-chatgpt-and-

could-it-repla ce-humans

Mason, K., Lindberg, K., Haenfling, C., Schori, A., Marsters, H., Read, D., & Borman, B. (2021). Social vulnerability indicators for flooding in Aotearoa New Zealand. *International Journal of Environmental Research and Public Health*, *18*(8). https://doi.org/10.3390/ijerph18083952

Militano, L., Arteaga, A., Toffetti, G., & Mitton, N. (2023). The Cloud-to-Edge-to-IoT Continuum as an Enabler for Search and Rescue Operations. *Future Internet*, *15*(2). https://doi.org/10.3390/fi15020055

Munawar, H. S., Mojtahedi, M., Hammad, A. W. A., Kouzani, A., & Mahmud, M. A. P. (2022). Disruptive technologies as a solution for disaster risk management: A review. *Science of the Total Environment*, *806*. https://doi.org/10.1016/j.scitotenv.2021.151351

Peek, L., Wachtendorf, T., & Meyer, M. A. (2021). Sociology of Disasters. *Handbooks of Sociology and Social Research*, 219–241. https://doi.org/10.1007/978-3-030-77712-8_11

Prasanna, R., Chandrakumar, C., Nandana, R., Holden, C., Punchihewa, A., Becker, J. S., Jeong, S., Liyanage, N., Ravishan, D., Sampath, R., & Tan, M. L. (2022). "Saving precious seconds"— A novel approach to implementing a low-cost earthquake early warning system with node-level detection and alert generation. *Informatics*, *9*(25), 1–32. https://doi.org/https://doi.org/10.3390/informatics9010025

Prasanna, R., Yang, L., King, M., & Huggins, T. J. (2017). Information systems architecture for fire emergency response. *Journal of Enterprise Information Management* , *30*(4), 605–624. https://doi.org/10.1108/JEIM-12-2015-0120

RNZ. (2023). *Cyclone Gabrielle updates: Death toll rises, half of uncontactable reports resolved* . RNZ. https://www.rnz.co.nz/news/national/484453/cyclone-gabrielle-updates-death-toll-rises-half-of-uncontactable-reports-resolved

UNDRR. (2017). *The Disaster Risk Reduction (DRR) Glossary*. UNDRR. https://www.undrr.org/terminology/disaster

UNISDR. (2009). *2009 UNISDR Terminology on Disaster Risk Reduction* . International Strategy for Disaster Reduction. https://www.undrr.org/publication/2009 -unisdr-terminology-disaster-risk-reduction

Vermiglio, C., Noto, G., Rodríguez Bolívar, M. P., & Zarone, V. (2022). Disaster management and emerging technologies: a performance-based perspective. *Meditari Accountancy Research*, *30*(4), 1093–1117. https://doi.org/10.1108/MEDAR-02-2021-1206

Witkowski, K., Remington, C. L., & Ganapati, N. E. (2021). Focus group research in disaster and emergency management. In J. Rivera (Ed.), *Disaster and Emergency Management Methods:* (Issue 1, pp. 123–141). Routledge. https://doi.org/10.4324/9780367823948 -10

Vision for Emergency Management: Conceptualizing Mission-Critical Ecosystems

Jaziar Radianti
Centre for Integrated Emergency Management
University of Agder, Norway
jaziar.radianti@uia.no

ABSTRACT

Emergency management is seldom viewed through the lens of enterprises/industries, which aim to safeguard their critical missions to deliver seamless services. This article advocates for an integrative perspective on mission-critical (eco)systems. In the digital transformation landscape, the traditional notions of "mission-critical systems," often associated with physical spaces like control facilities or hardware-software crucial for business operations, are challenged. Based on document reviews, this article argues for a holistic comprehension of an organization's mission-critical ecosystems, considering internal and external factors across physical, digital, and human aspects. The article uses renewable energy and e-health as cases. The overarching goal is to expand the research scope in emergency management by examining current concerns and current trends in emergency management domain, combining preparedness, resilience, and cybersecurity. This comprehensive approach aims to enhance our ability to navigate and respond effectively to emergencies in an increasingly complex and interconnected world.

Keywords

Mission-critical ecosystems, digital resilience, risks, cybersecurity

INTRODUCTION

Numerous studies in the emergency management domain often focus on cooperation, coordination, and collaboration among public responder agencies to save lives and protect properties. However, many organizations including private companies and industries are also vulnerable to physical and digital threats and often dependent on external actors to conduct their operations. Business entities in society play critical roles as pillars for economic resilience. Some businesses are essential in running critical services such as companies involved in supplying energy, health, and care services. In operating the businesses, there are always specific entities of private or public organizations encompass persons, information systems, and digital infrastructures that are critical, and should continuously be protected to avoid operational disruptions, i.e., mission critical functions. The sources of interruptions to mission-critical functions can originate from either inadvertently or advertently made disasters. The impacts can be harmful, ranging from the loss of reputation, trust, monetary assets to the loss of human lives. These critical units, internal infrastructure, actors, or services in this article are referred to as "mission-critical ecosystems."

Rooted from the ecology science (Mars et al., 2012), the concept of ecosystems itself has been metaphorically applied across various scientific areas (Pickett & Cadenasso, 2002), including organizational ecosystems and management (Mars et al., 2012; Pavlikakis & Tsihrintzis, 2000), or even digital ecosystem metaphor (Krivý, 2023) that can cover the whole planet. Scholars acknowledge the complexity of ecosystem concept, which is subjected to continuous change in space and time, characterized by interconnectedness, dynamics, and scales—at which ecosystem processes operate, with human playing a significant role (Pavlikakis & Tsihrintzis, 2000). Boundaries of ecosystems exist but they may be defined arbitrarily because of spatial and temporal changes of the ecosystem. The open and dynamic character of the ecosystems does not support rigid guidelines for boundary delineation

WiP Paper – Visions for Future Crisis Management
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

(Pavlikakis & Tsihrintzis, 2000). In this article, mission critical ecosystems (MCEs) are defined as an integration, and interconnectedness between the critical *internal arrangement* of organizations (actors, knowledge, digital platforms, critical facilities, supply chain) with *external stressors* that pressing organizations to adapt, respond or change to fulfill their goals (e.g., protect essential service functions or core businesses) which sometimes require interactions with external stakeholders. Thus, critical infrastructure can be one of the components of mission critical ecosystems, while critical services can mean the core businesses of industries fall within societal critical functions or they can mean any critical functions in organizations that vital for seamless business operations.

Indeed, there are substantial body of literature on business continuity management (BCM) and continuity of operations (COOP) planning (Cerullo & Cerullo, 2004) which addresses the maintenance of an organization's mission-critical functions during disruptions to critical infrastructure caused by various physical, cyber, public health and other hazards. This focus is also a part of governmental emergency management efforts as evidenced by publications from the US Federal Emergency Management Agency - FEMA (2018; 2020) and Cybersecurity and Infrastructure Security Agency - CISA (2024), which tackle both physical and cyber threats. Notably, as a managerial function concerned with coordinating "whole community" efforts, emergency management in these publications underscore the necessity of involving industry and community organizations in mitigation and preparedness activities to uphold essential services during crises.

COOP literature offers practical guidelines to cope with the disturbances, but they do not fully address what counted as MCEs. Moreover, MCEs are becoming blurry with recent digital transformations when industries gradually rely heavily on internet of things, cyber-physical (social) systems, internet of service and smart factory. Reliance on "interconnectedness" between devices, applications, sectors, information/ digital/ physical infrastructures are a key in the so-called industry 4.0 premise. Moreover, such "smart arrangement" often requires elevated needs of hardware, software, and services to make mission-critical infrastructure to function in a more efficient manner through outsourcing to the third parties, and sometimes go beyond the national boundaries. Thus, physical, digital, and service supply chain are unavoidably a part of the MCEs. Todays´ ubiquitous connectivity, the use of Internet of things and cloud services, digitalization of critical infrastructure including health services, and other digital trends, have triggered the needs for understanding such landscape from the organizational viewpoints, and the idea of the digital disasters (Hansen & Nissenbaum, 2009; Van der Meulen, 2013) and digital resilience (Boh et al., 2023; Magutshwa & Radianti, 2022).

Digital resilience often includes building resilience of entities that are vulnerable to cyberthreats including human, technologies, and digital infrastructures. Thus, cybersecurity aspects are inseparable from societal security. The Covid-19 pandemic can be considered as a health disaster, that further triggers emergency management, and the need for understanding cybersecurity better, because solution services are delivered digitally (Magutshwa, 2022; Magutshwa & Radianti, 2022). Another example, many countries such as Norway are in the energy transition from usage of fossil-based energy fuel to renewable energy and battery technology. There have been intensive political and scientific discussions for building more energy sources such as offshore and onshore wind power (Dahl et al., 2022), moving toward battery sources for ground, air, and sea transports (Figenbaum, 2023; Schulz & Rode, 2022; Thorne, Hovi, et al., 2021; Veisten et al., 2024).

The significant shift in Europe, particularly in Norway, aims to prepare for future scenarios with limited fossil fuels by exploring new forms of renewable energy beyond hydroelectric power (Skjærseth & Rosendal, 2023), as one of the oldest and largest sources of renewable energy. While various renewable energy sources exist (solar, geothermal, and bioenergy), Norway leans towards among other things offshore and onshore wind power, aligning with higher prominence in the political agenda (NOU, 2023a). Additionally, battery power emerges as a pivotal energy storage, especially with the increasing adoption of electric vehicles in Norway. For instance, in 2023, 83.4% of new cars sold in the country were electric (source: www.elbil.no), equipped with 40-50 kilowatt-hours (kWh) batteries (Innovasjon-Norge, 2023). City buses as critical transport infrastructures are gradually adopting batteries with an average capacity of 125 kWh.

The infrastructures controlling these new sources of renewable energy are also changed, and more digitalized. From a societal perspective, companies provide critical services facing unknown risks, challenges, societal conflicts, and potential disasters, including digital disasters, which require new understanding of emergency management. Again, the MCE is a way to entangle interrelated aspects of future emergency management involving new digital-related challenges and digital disasters.

Moreover, the increasing use of digital health services, as evidenced in the e-health case example (Helsedirektoratet, 2023), will require a robust communication infrastructure. This transformation undoubtedly elevates the demand for a more seamless electricity supply (Thorne, Aguilar Lopez, et al., 2021). This change is intricately linked to the current green ambitions in Europe towards achieving net-zero emissions. This overarching

goal is encapsulated in the vision to "Make Europe the first climate-neutral continent by 2050" (EU-Commission, 2023).

This paper argues on the need of moving our understanding from mission-critical *environment* and *systems* to *mission-critical ecosystems*. In the industrial context, numerous layers of entities, units, and environments interact to influence MCEs, which are essential for delivering sustainable services. I propose the following research questions:

- RQ1 How to define a framework for studying mission critical ecosystems?
- RQ2 How to apply the mission critical ecosystem framework to derive emerging issues within critical services?

We need to gain better understanding of what are counted as MCEs. This paper aims at contributing to research agendas to define and expand the idea of MCE as one of potential future direction for emergency management research. The cases and examples were taken from European setting especially Norwegian example, with two selected critical services, i.e., the e-health service and energy sectors.

This paper´s contributions are threefold. *First*, it proposes a definition of the MCE concept that guides this study to find the research gaps in existing literature related to business continuity/ emergency management. *Second*, the MCE concept is essentially examining the wider set of components not only infrastructural interdependency, but also issues such as green shift, new emergency management knowledge and competencies due to changing threat landscape. Current approaches to business continuity/emergency management planning and preparedness efforts are likely to overlook these dependencies, rendering essential services vulnerable during crises. *Third*, the paper proposes an integrated framework i.e., MCEs and applies to examine strategic documents to identify future new challenges that are worth research.

LITERATURE REVIEW

In this section two interrelated concepts are elaborated based on the literature review, i.e., mission critical systems and business continuity planning. Finally, the section summarizes research gaps and highlights this paper's position.

Mission-Critical Systems in the Literature

The complexities described earlier trigger the need to build a new understanding of "mission-critical (eco)systems" that are now becoming blurry. Traditionally, mission-critical systems are often associated with the critical physical infrastructure and facilities, and often are associated with the 24/7 environments such as operation centers, network equipment rooms, standby emergency power, business continuity and technology recovery room, business operation control room, command center—to name few (Curtis, 2021; Hebert et al., 2018). There are many specific rooms and areas within facilities in today's ever-changing environment. In this book, Curtis (2021), defines the Critical Environment (CE) as the "*physical space and the systems within a facility that are uniquely configured, sized and dedicated to supporting specific critical business operations as defined by the user*". Before 2010, mission-critical systems often were associated with Mission-critical Computer Resources (MCCR) such as hardware/software/firmware (Austin & Larkey, 1992) and critical servers (Johnson, 1999).

Table 1 summarizes various definitions used in the literature, to broaden our understanding, how the concept has been applied in contexts beyond "critical facilities" and "critical systems."

Table 1 Overview of the mission critical concepts derived from literature.

Main Concepts	Definition
Systems or Process (Bernardini et al., 2013; Magutshwa & Radianti, 2021; Weger et al., 2023)	**A system or process** is one in which a failure or interruption comes with intolerable operational or human cost. The system is **essential to the survival of an organization** and **will adversely affect society when it fails.**
Systems health state; socio-technical interaction (Bernardini et al., 2013)	A 'mission-critical system' is a system whose '**health state**' is crucial to the successful completion of an organization's core operations. Many mission-critical systems can be characterized as **complex socio-technical systems**, that depend on the interactions among human and social factors and technical infrastructures.
Criticality scale (Microsoft, 2023), *Approaches, and methods* (LeSaint et al., 2015)	**A criticality scale** that covers significant financial cost (business-critical) or human cost (safety-critical) **associated with unavailability or underperformance**. As an approach, it refers to methods for assessing vulnerability in mission-critical systems.

Table 2 Overview of the mission critical concepts derived from literature.

Main Concepts	Definition
Function or Critical Services (Zuo, 2013)	**A function whose failure would lead to catastrophic consequences** that would place public security at risk. Mission-critical systems provide **vital services and must be reliable and dependable** to withstand malicious attacks and system failures (e.g., the electric power grid, telecommunications networks, healthcare systems, water management systems), hardware and software failures, operator errors, power outages, environmental disasters, and attacks by adversaries
Hardware, software, applications, communication facility, IoT (Ali & Ware, 2021; Austin & Larkey, 1992; Cao et al., 2021; Johnson, 1999; Weger et al., 2023)	**The hardware, software, and communication facilities**, that allow mission-critical users to communicate with each-other and liaise with command centres securely for providing mission-critical services, wherever and whenever the services are needed. It includes applications that critical, e.g. to support Internet of Things (IoT) that essential for autonomous driving, factory automation and tele-surgery.
Business Ecosystems (Mattos et al., 2020)	**Continuous process in a business-to-business (B2B**) mission-critical systems: customer and users can be different customer can subscribe for their users, or they become users in itself; Ownership products and data issues
Leadership (Yip et al., 2009)	**Leadership as a mission-critical:** Horizontal and vertical boundaries, requirements for groups to work across boundaries. Traditionally, managers learn to manage vertically — to work with senior colleagues and downward with direct reports. In today's interconnected world, it is critical for managers to work effectively across functions, locations, and with external stakeholders
Innovations in public agencies (Breznitz et al., 2018)	**Enhancing innovations**: mission-critical systems have been linked into the need for tools and structure of successful innovation agencies to be one of the following organizational typologies: upgrader, productivity facilitator, state disruptors, transformation enabler
Organizational capacity (Hannah et al., 2009)	One or more extreme events are occurring or are likely to occur that may **exceed the organization's capacity** to prevent and result in an extensive and intolerable magnitude of physical, psychological, or material consequences to – or in close physical or psycho-social proximity – to organization.

The usage of the "mission-critical" concept in Table 1 shows the varieties of this concept applications, which are then aggregated to define the component of MCEs. The learning points from this review are:

- "Mission-critical" (systems, processes) are often treated as a singular topic, either specifically discussing the concept within a particular application area or in a very general sense to refer to important activities.
- The term refers to both the technical development of mission-critical software, apps, and communication, and to the critical spaces, or facilities.
- In the organizational context, the term is often associated with risks, the sustainability of business operations, structure, leadership, and the organizational capacity to cope with stress. Frequently, the term is related to the concept of "cascading effects" resulting from the failure of one critical service, impacting other sectors and business units.

Literature on Business Continuity Planning (BCP)

A related perspective in the literature that addresses the maintenance of business operations is CP framework. BCP entails preplanning disruptive events (Reid, 2021), establishing a plan to counter mitigate risks and minimize the impact of a crisis while reducing the time needed to restore conditions to normal operations (Cerullo & Cerullo, 2004). The aim of BCP is to prepare for, provide to, and maintain control and capabilities for managing an organization's overall ability to prevent operational discontinuity during disruption (Mukherjee et al., 2020).

Reid (2021) points out that BCP is analogous to Continuity of Operations Planning (COOP), with the distinction that BCP is often applicable for private sectors, while COOP typically refers to continuity of government or public sectors, which has been accepted as a part of the emergency preparedness landscape since 1955 in the US (Rucks et al., 2011), and continuously updated (FEMA, 2020). Rucks et al. (2011) that suggest templates developed for emergency management agencies during planning stages and their utility during activations or response phases. FEMA also provides practical examples of templates for continuity plans (FEMA, 2020). Thus, BCP is often manifested as a playbook designated to guide response in the event of disruptions. Specifically, COOP illuminates the efforts within individual organizations to ensure that essential functions continue during disruption of normal operations (FEMA, 2018, 2021). Typical instruments to identify critical functions include conducting Business

Impact Analysis (BIA), assessing associated risks, and formulating mitigation plans (Cerullo & Cerullo, 2004; FEMA, 2019) as well as creating Disaster Contingency Recovery Plan (DCRP) to outline procedures to follow when a disaster occurs (Cerullo & Cerullo, 2004).

Inspired by Covid-19 crisis, Grace et al. (2023) introduce continuity planning for cyber-physical infrastructure to maintain essential functions at existing or alternative workplaces thorough the crisis. The article discusses the interdependence of cyber-physical-social infrastructure in continuity planning for public health emergencies and suggests the need to expand business process/impact analysis to recognize dependencies and impacts on critical infrastructure that occur when other organizations' essential functions are impacted. Grace et al. (2023) particularly highlight the need for multiple-level redundancies to maintain continuity of operations and achieve organizational resilience.

Summary and Research Gaps

In the business and industrial context, highlighting the urgency of emergency management is clearer when using concepts like BCP or "mission-critical ecosystems," as any threats to digital and physical infrastructure, critical functions, or lacking critical teams and leadership are immediately recognized as crises requiring response. Existing studies, however, appear to focus on a limited set of infrastructures and not the expansive set of inter-organizational infrastructure components—ecosystems—that organizations depend on to maintain operations during a crisis. One of the gaps in the BCP framework is its inadequateness in addressing emerging risks such as new cyberthreats, climate change impacts and zero emission requirements assigned to organizations. This includes considerations such as prioritizing human centric approaches over reliance on automation, integrating climate risk management for sustainability, and incorporating social dynamics, as suggested by Industry 5.0.

Moreover, there is a gap where the mission-critical (eco)systems are not portrayed holistically, by considering all the ecosystems, and allowing the organizations to assess their weaknesses in a more thorough way. This article proposes that "Mission critical ecosystems encompass core technologies, physical infrastructures, the individuals, and teams who interact with the technologies, organizational and supply-chain environments that support critical (business) operations. A failure in one or more of these factors within this ecosystem can result in undesirable cascading risks and consequences for both business and society."

METHODOLOGY

Approach: Multi-Vocal Literature Review (MLR) Approach

This study is designed as qualitative research involves exploring research questions through collecting and analysing data using typical qualitative data collection methods. The study employed ***multi-vocal literature review (MLR) approach,*** especially because we want to include overarching themes laid down by strategic documents related to selected cases (e-health and energy). "MLR comprises all accessible writings on a common, often contemporary topic. The writings embody the views or voices of diverse sets of authors (academics, practitioners, journalists, policy canters, independent research and development firms, and others).

They reflect different purposes, perspectives, and information bases. They address various aspects of the topic and incorporate different research or non-research logics" (Garousi et al., 2019, p. 102).

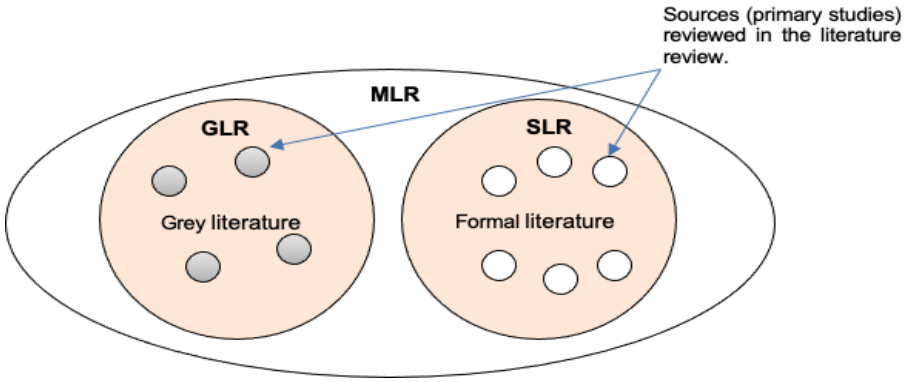


Figure 1 Venn diagram showing the relationship of SLR, GLR and MLR studies (Adapted from Garousi et al, 2019).

Data Collection Method from Academic and Grey Literature

This study uses documents as the main data sources, which then be coded and analysed to find patterns and themes, i.e. scientific literature, and grey literature (GL). The usage of grey literature in research have been discussed by scholars. The 1997 Luxembourg Convention defines the term as "*literature that is produced on all levels of government, academics, business, and industry in print and electronic formats, but which is not controlled by commercial publishers i.e., where publishing is not the primary activity of the producing body*" (Garousi et al., 2019; Kamei et al., 2021). Such literature is not obtainable through traditional publishing channels that typically require peer-review processes prior to publication to control publishable materials. Garousi et al. (2019) suggest three levels of GL. **The first tier** encompasses books, magazines, government reports and white papers with high outlet controls and high credibility such as reports, books, magazines government reports and white papers. **The second tier** has moderate outlet control and credibility such as annual reports, news articles, presentations, videos, question-answer sites, Wiki articles and so on. **The third tier** encompasses e.g., blogs, emails and tweets as low outlet control and credibility. Scholars argues on the importance of GL as it has the potential to understand the contextual contemporary information and provides valuable insights for research in fields where there are still received limited scholarly attention (Garousi et al., 2019).

The overall procedures conducted in this study is captured in Figure 2. In this study, scientific literature was used to map the components of "mission-critical (eco)systems". The search term "mission-critical" were applied in Scopus database and Google Scholar. Highly technical-oriented papers were filtered out. The selection focused on articles that tailor mission-critical term with organizational context. This procedure help deriving the mission critical ecosystem framework that integrates its component into a single framework (Figure 3). On the GL part, this study benefited from the government strategic documents publicly accessible, published between 2019-2023 by Norwegian Government or agencies that touch upon energy sectors, emergency preparedness, health preparedness and the green shift. Using the framework derived from the literature in Figure 3, I extracted emerging themes from the identified government reports that serves as vision for emergency management.

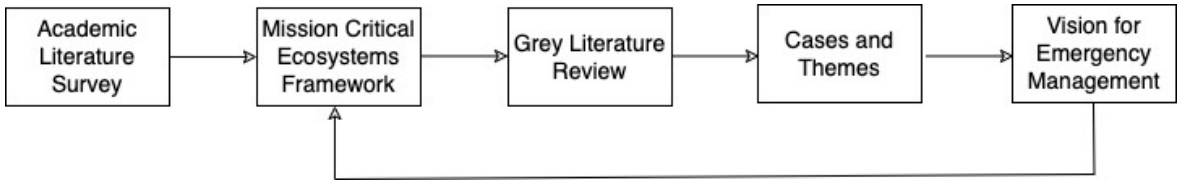


Figure 2 Procedure for deriving vision for emergency management

Data Analysis Method

The documents were coded thematically using a qualitative software Nvivo with simple criteria: renewable energy and health and societal security topics. However, whenever a new concept comes, it was treated as a new sub-theme during the analysis, which then was mapped into the broader theme. This framework was then used to map the government documents and strategies that dealing with more concrete e-Health and energy cases and to derive visions for future emergency management.

RESULTS

Figure 3 depicts a simplified MCEs framework derived from scientific literature, which subsequently utilized to map emerging themes within selected application areas and beyond. Figure 3(A) lists sources of instability both internally (on the right side) and externally (on the left side). The core circle captures various necessary applications, software, hardware, and various technologies and digital infrastructures needed to sustain the organizational operations. The second inner circle represents the "space", where these supporting technologies reside, while the third inner circle (mission-critical users) comprises the core teams and human resources interacting with the technologies, ensuring seamless operation of the organizations. The fourth inner circle represents the criticality of organizational layer where the leadership, organizational cultures (safety, security, and risk cultures). The outermost circle portrays the supply-chain of goods (e.g., raw materials, equipment, critical components of industries), services (consultancies, software as a service, expertise for hire), and digital supply chain (e.g., digital technologies, data analytics, predictive analytics, Internet of things (IoT)-based services). While in Figure 3A the whole landscape of the MCEs seems disjointed, in reality there are dependencies among these factors as has been simplified in Figure 3B.

Externally (left side of Figure 3), there are known and unknown factors that affect businesses´ resilience such as:

- extreme events (e.g., snowstorm, floods, terrorist actions, landslide, pandemic)
- risks, threats and cyberthreats (e.g., black-swan events, cyberattacks, cybercrime and extortion)
- new demands for customers (e.g., preferences, desired services, sophistication levels, platforms)
- contemporary trends (e.g., industry 4.0, industry 5.0, sustainability, green deals demand, new competencies)
- new national and international strategies (e.g., stronger privacy laws, new highlights in governments´ regulations, artificial intelligence act, and so on)

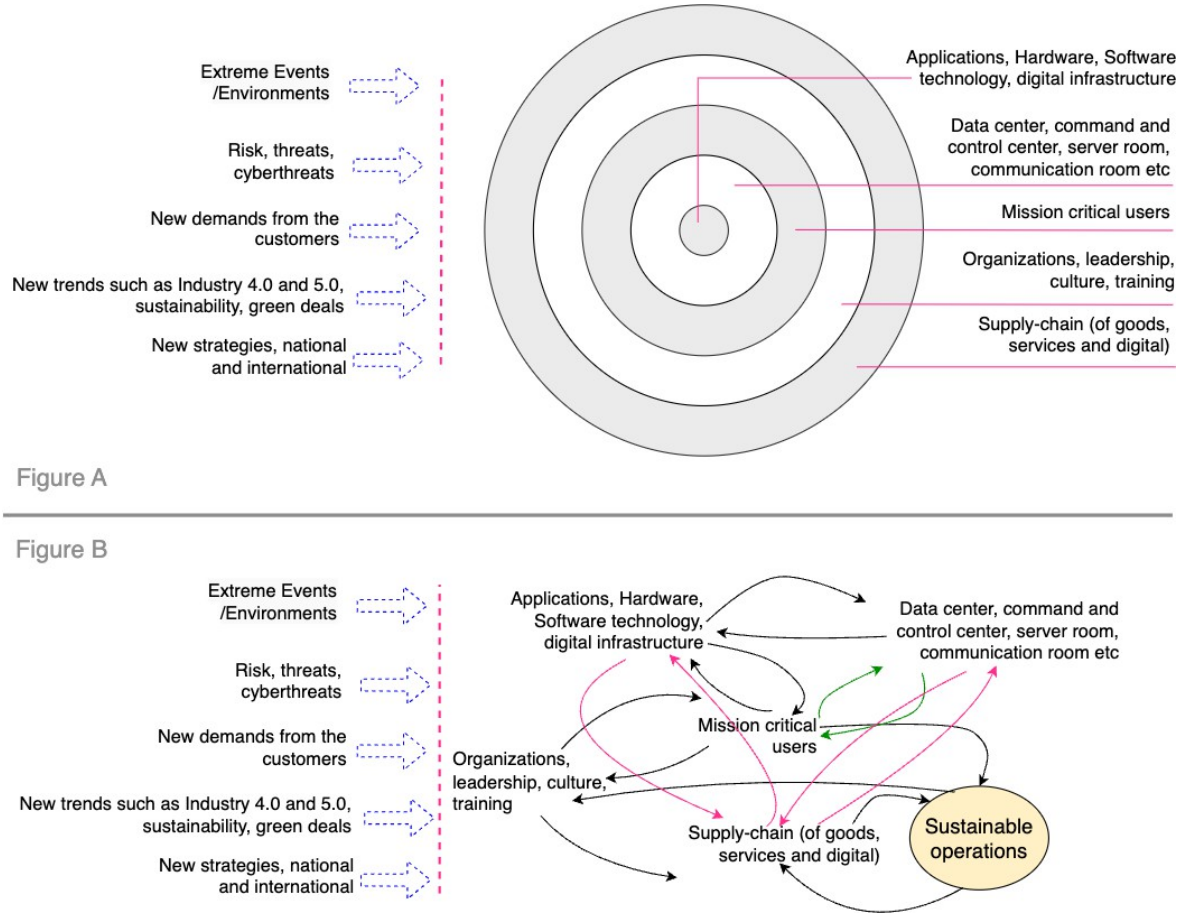


Figure 3 A Framework to untangle the mission-critical ecosystems.

We applied this MCEs as thematic analysis framework into two selected cases, i.e., challenges in renewable energy and e-health sectors that affect societal security. The following trends are identified and reemphasized:

Applying the MCE Framework for the e-Health Case

As our focus is to identify the e-health themes, the following issues within the MCEs framework were identified:

- **Digital and critical infrastructure**: Themes such as digital infrastructure, information security, managing security breaches, secure communication, security-privacy consequences, risk assessment and management, and secure IT operation with access control have repeatedly emphasized (DEH, 2023; Meld.St.5, 2023; NORMEN, 2018, 2022a, 2022b, 2023). The ICT solution is a backbone of the eHealth service. Organizations should capable of assessing the risks and consequences regarding the unavailability of the service (NORMEN, 2021).
- **Facility:** The inclusion of e-health initiatives in health preparedness aims to enhance monitoring, detection, and analysis of e-health digital threats and vulnerabilities, the relevance of Computer Emergency Response Teams (CERTs), and required facilities for conducting these activities (Meld.St.5, 2023).
- **Critical users:** There are lacking competence and knowledge in some sectors including basic research in the development of pharmaceutical, diagnostics and vaccines, medical technical equipment, and digital health.

The health sector should be prepared against crisis and catastrophes (*Veikart Helsenæringen*, 2023), and ICT security competence to understand threats and risk control (Meld.St.5, 2023).

- **Organization, leadership, and culture:** The themes touch upon the issues of strengthening system for health preparedness, resilient health preparedness, and understanding of risk and vulnerability. Health organizations should cooperate on societal security based on systematic risk management. Prevention and preparedness must be prioritized (Meld.St.5, 2023). In eHealth, the organizations should comply with the requirements for information security (NORMEN, 2022a).
- **Supply chain**: Sometimes health organizations act as data controller by collecting health data, which may be processed further by the third-party data processors on behalf of the data controller. Data controller organizations sometimes act both as data controller and data processor. In other instances, suppliers come as data processors in the cloud services or operational services, maintenance (physical service) and remote access, or even outsourcing the ICT functions. These suppliers should adhere to the existing requirements to avoid potential data breaches (NORMEN, 2022a)
- **External issues:** Examples of themes include feared threats and risk scenarios involving digital infrastructure in e-health including human errors, accidents, technical failures, and extreme weather are highlighted. The risks are annually assessed by relevant national agencies such as policy security service, intelligence services, national security authority and Health-CERTs. Relevant risks, threats and vulnerabilities that will harm the availability, confidentiality and integrity of personal information should be understood (NORMEN, 2022b). There are also concerns on software dependency, dependencies related to the supply chain and digital infrastructure, as well as the development of artificial intelligence, and its consequences (Meld.St.5, 2023).

Applying the MCE Framework for Power Sector Scenario: "New" Renewable Energy and Battery

The themes derived from energy case are coded and aligned the MCEs framework, for instance, automation and digitalization, which significantly contribute to increased power consumption (Motvind-Norge, 2019), alongside the electrification of the transport systems. The following themes were identified:

- **Technology and digital and critical infrastructure**: The energy storage system such as battery is a key for keeping excessive energy production originated from use of renewable energy such as wind power (Motvind-Norge, 2019). Such energy source development has changed the risk picture from the societal security perspective (NOU, 2023b). The energy white-paper document (NOU, 2023a) indicates the emergency management work should include the understanding of the new incident risks caused by such power sources and how to manage them (NOU, 2023b).
- **Facility:** Many energy companies active promoting the energy-intensive data center and the energy usage will increase, thus this development is a part of the calculated future energy needs (NOU, 2023a)
- **Critical users**: The theme stresses the shortage of personnel with practical risk management skills, hindering the effective implementation of risk analysis and measures in practice (Johnsen & Lysgaard, 2022). This knowledge gap can impede the procurement of ICT systems, particularly in assessing their security risks. There is a pressing need for clearer roles for policymakers and other stakeholders to enhance the resilience of critical infrastructure and systems (NOU, 2023a)
- **Organization, leadership, and culture**: the themes point out the need for a better coordination and cooperation among relevant organizations such as Power supply Emergency Preparedness Organizations (KBO). Culture for sharing information is also need improvement (NOU, 2023a, 2023b).
- **Physical and digital supply chain:** the transition to e.g., wind power or battery, triggers higher demand for raw materials, minerals, and metals. Thus, industries are more dependent on a number of goods such as aluminum, cobalt, copper, steel, zinc and lithium (Motvind-Norge, 2019). The supply chain also involves complex arrangement of suppliers involved in infrastructure construction. The digital supply chain receives high attention due to its role as a gateway for cyberattacks, with statistics indicating a rising trend in attacks among suppliers, leading to digital disasters. It encompasses the delivery structure between businesses, involving digital services, software, or hardware in each delivery (DSB, 2020). Thus, ICT security risk management of the supply chain is especially important in many sectors, especially energy (Johnsen & Lysgaard, 2022). Understanding the supply chain is becoming relevant and crucial to safeguard mission-critical core of the business.
- **External issues**: There is high fluctuation of electricity supply from "new" energy such as wind power. In certain European countries, dual wind power sources, offshore and onshore, are common, but they present challenges such as construction and funding. The potential of onshore energy for power balance is recognized, yet barriers have caused delays, due to environmental concerns and local conflicts, such as community acceptance of onshore wind facilities. Harsh environments can heighten the risk of infrastructure damage

(NOU, 2023a). Likewise, on the offshore part, the risks of marine ecosystem damage and other ocean-related interests can be affected negatively. These are an example as well how the mission-critical business core can affect societal security, especially with the societal conflict risks, and vulnerable physical infrastructure.

Societal Security and Cybersecurity in relation to eHealth and Energy Scenario

A strategic document called "Risk 2023" advices businesses to have better visions on security work in a larger context, and take measures to protect themselves against espionage, sabotage, terror, and complex threats (NSM, 2023). Kvadsheim et al. (2023) pinpoint that the security quality is closely connected the preparedness. Energy-related organizations need to establish preparedness plan (how to respond), structure (role and responsibilities), training (to test if the preparedness plan works), evaluation and learning. The Norwegian *Total-preparedness* strategy justifies the preparedness and management in energy supply and recommends measures to handle unwanted threats (NOU, 2023a), including digital security and preparedness for digital response (NOU, 2023b).

The *total-preparedness* strategy covers also health preparedness, aimed at safeguarding lives and well-being by ensuring continual provision of medical treatment, nursing, and care for individuals affected by crises, including major natural disasters, terrorist attacks, war, and pandemics and typically causing death or severely injured victims (NOU, 2023b). This health preparedness highlights the importance of services in emergency medicine, emergency room, municipal health service and specialist health services. Swine-flu pandemic 2009 and Covid-19 pandemic are examples of the importance of good infection controls as public health measure. Overall, the *total-preparedness* strategy underlines improved risk analysis and management for better health preparedness, including the protection of critical infrastructure to maintain seamless electronic communication services in crisis events. Battery has been mentioned as an important energy storage, to back up the energy sources for critical infrastructure during the power outage.

DISCUSSIONS AND REFLECTIONS

The identified trends or issues highlighted in the previous sections indicate that organizations are often vulnerable, facing both new and interrelated threats that may go unnoticed without a comprehensive analysis of the overall ecosystems. The listed trends in the previous section, derived mainly from the Norwegian experience, specifically in e-health and energy sectors, may not align with other countries' experiences. For example, differences in the adoption of electric cars, levels of digitalization for critical infrastructure systems, reliance on electricity in daily life, and dependence on renewable energy sources vary among nations. In fact, Norwegians build their preparedness system based on trust, a societal value that highly appreciated in their society, may not be applicable to countries with more fragmented social structures and higher level of distrust. However, certain general topics consistently emerge, and generally applicable in building emergency management:

- The strategic documents emphasize the significance of ***the preparedness stage*** in emergency management. A well-crafted preparedness plan, outlining response procedures, organizational structure, role delineation, training protocols, and continuous evaluation, is crucial for an effective crisis response. Systematic research on building excellent preparedness plans across diverse and emerging scenarios could be a relevant research agenda.
- The ability to conduct thorough ***risk assessments*** is pivotal in creating realistic preparedness plans that account for new threats and are regularly updated. Research on risk analysis, risk perceptions, systemic risks analysis, intertwined risks, and cascading risks within the emergency management and response domain could contribute significantly to the field.
- Examining ***organizational ecosystems*** is crucial in establishing emergency plans and conducting risk analyses. The proposed mission-critical framework, focusing on each component or layer, aids in understanding threats and defining preparedness plans.
- ***Digital transformation*** is inevitable in various sectors, including emergency management. However, limited studies exist on digitalization within this domain, including potential "digital disasters." Research on the implications of digital disasters and their relevance to emergency management is needed.
- The increasing connection between ***cybersecurity and emergency management*** is noteworthy. Understanding digital risks, security, and privacy protection becomes integral to post-digital transformation. Research should explore events that resemble cybersecurity incidents but trigger emergency management responses, such as a cyberattack on a transport control system causing citywide traffic chaos.
- Resources with skills in risk analysis, cybersecurity, and ICT supply chain security are limited. Research on ***human resources in the digitalization era*** within emergency management, focusing on capabilities, competencies, and training ideas to address gaps, is essential.

To sum up, this study provides valuable lessons that can guide efforts to derive new research areas in the emergency management domain that put business and industries into the picture, particularly when examining individual scenarios like those are relevant for the e-health and energy as critical services. Moreover, to repeat that we have proposed the definition MCEs as an integration, and interconnectedness between the critical *internal arrangement* of organizations (actors, knowledge, digital platforms, critical facilities, supply chain) with *external stressors* that pressing organizations to adapt, respond or change to fulfill their goals (e.g., protect essential service functions, core businesses) which sometimes require interactions with external stakeholders. By applying this framework during the analysis stage, several pressing issues on each category of component in mission critical ecosystems have been identified. The implication for research among other things is how to identify detailed components of mission critical ecosystems, and how to detail the green transition within industrial mission critical ecosystems. There is a newer model in the BCP literature such as Eco-centric BCP that build on four pillars: 1) reduction in emission from industry; 2) sustainable use and reuse of ecosystem services; 3) legislation-and implementation, and 4) sustainable energy usage and consumption. In the Eco-centric BCP Model, environmental aspects are highly dominant and translated as e.g., environmental norms, waste production and reduction, emission treatment, certification for sustainable indexing, circular economy, and emission management. This can be explored as further study encompassing their relevance for mission critical ecosystems, and how organizations move forward with the eco-centric BCP (Mukherjee et al., 2020).

CONCLUSION AND FUTURE WORK

This article argues for the necessity of adopting a less siloed approach to emergency management. As technology plays an increasingly integral role in emergency response, the distinction between emergency management and cybersecurity becomes more challenging. Various technologies are now utilized in emergency response activities, encompassing tasks such as data collection, exchange, and storage. The proposed mission-critical framework serves to map and comprehend the emergency management landscape from an organizational standpoint. This study relies on a single-country case and a few industrial or business scenarios, however, it sheds light on diverse and promising research directions. Moving forward, it is essential to explore additional scenarios, examine deeper into case studies and leverage the mission-critical ecosystem as a framework for enhancing emergency management preparedness and response.

REFERENCES

Ali, A., & Ware, A. (2021). Effective Performance Metrics for Multimedia Mission-critical Communication Systems. *Annals of Emerging Technologies in Computing (AETiC)*, *5*(2), 1-14.

Austin, R., & Larkey, P. (1992). The unintended consequences of micromanagement: the case of procuring mission critical computer resources. *Policy Sciences*, 3-28.

Bernardini, G., Paganelli, F., Manetti, M., Fantechi, A., & Iadanza, E. (2013). SYRMA: a tool for a system approach to risk management in mission critical systems. *International Journal of Business Information Systems*, *13*(1), 21-44.

Boh, W., Constantinides, P., Padmanabhan, B., & Viswanathan, S. (2023). Building digital resilience against major shocks. *MIS Quarterly*, *47*(1), 343-360.

Breznitz, D., Ornston, D., & Samford, S. (2018). Mission critical: the ends, means, and design of innovation agencies. *Industrial and Corporate Change*, *27*(5), 883-896.

Cao, J., Zhao, J., Zhu, X., Jiang, Y., & Wei, Z. (2021). Toward a green secure relay system for mission-critical IoT: Hybrid duplex relay selection and resource allocation in the finite block length regime. *IEEE Transactions on Green Communications and Networking*, *5*(4), 1869-1879.

Cerullo, V., & Cerullo, M. J. (2004). Business Continuity Planning: A Comprehensive Approach. *Information Systems Management*, *21*(3), 70-78. https://doi.org/10.1201/1078/44432.21.3.20040601/82480.11

CISA. (2024). *Emergency Services Sector Continuity Planning Suite.* . USA: Cybersecurity and Infrastructure Security Agency Retrieved from https://www.cisa.gov/emergency-services-sector-continuity-planning-suite

Curtis, P. M. (2021). *Maintaining Mission Critical Systems in a 24/7 Environment*. John Wiley & Sons, Inc.

Dahl, I. R., Tveiten, B. W., & Cowan, E. (2022). The Case for Policy in Developing Offshore Wind: Lessons from Norway. *Energies*, *15*(4), 1569.

DEH. (2023). *National e-Health Strategy*. Norway: Directorate of e-Health

DSB. (2020). *Risikostyring i digitale verdikjeder*. Norway: Direktoratet for samfunnsikkerhet og beredskap

A Green Deal Industrial Plan for the Net-Zero Age, (2023).

FEMA. (2018). *Continuity Guidance Circular*. USA: Federal Emergecy Management Agency - National Continuity Programs Retrieved from https://www.fema.gov/sites/default/files/2020-07/Continuity-Guidance-Circular_031218.pdf

FEMA. (2019). *Business Process Analysis and Business Impact Analysis User Guide*. USA: Federal Emergency Management Agency Retrieved from https://www.fema.gov/sites/default/files/2020-07/fema_BPA-BIA-Users-Guide_070119.pdf

FEMA. (2020). *Continuity of Operations Plan Template for Federal Departments and Agencies*. Washington DC, USA: Federal Emergecy Management Agency - National Continuity Programs Retrieved from https://www.fema.gov/sites/default/files/2020-10/fema_planning-template-federal-departments-agencies_october-2020_0.pdf

FEMA. (2021). *Developing and Maintaining Emergency Operations Plans Comprehensive Preparedness Guide (CPG) 101 Version 3.0*. USA: Federal Emergecy Management Agency - National Continuity Programs Retrieved from https://www.fema.gov/sites/default/files/2020-07/Continuity-Guidance-Circular_031218.pdf

Figenbaum, E. (2023). The contribution of research and knowledge accumulation in the development of the Norwegian battery electric vehicle market. *Transportation Research Procedia*, *72*, 4127-4134.

Garousi, V., Felderer, M., & Mäntylä, M. V. (2019). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and software technology*, *106*, 101-121.

Grace, R., Gautam, S., & Tapia, A. (2023). Continuity planning for public health crises: Designing workplace redundancies for organizational resilience. *Journal of emergency management (Weston, Mass.)*, *21*(6), 523-537.

Hannah, S. T., Uhl-Bien, M., Avolio, B. J., & Cavarretta, F. L. (2009). A framework for examining leadership in extreme contexts. *The Leadership Quarterly*, *20*(6), 897-919.

Hansen, L., & Nissenbaum, H. (2009). Digital disaster, cyber security, and the Copenhagen School. *International studies quarterly*, *53*(4), 1155-1175.

Hebert, P., Clare, G., Jayadas, A., & Balasubramanian, M. (2018). Tunable White Light System for Mission-Critical Control Room and Anti-Fatigue Room for Shift Workers: A Case Study. 2018 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC),

Innovasjon-Norge. (2023). Tall om batterier.

Johnsen, I.-K., & Lysgaard, V. (2022). *Risikostyring av IKT-sikkerhet i leverandørkjeder* Oslo, Norway: Norges vassdrags- og energidirektorat

Johnson, D. (1999). Secure Access to Mission-Critical Applications. *Inf. Secur. J. A Glob. Perspect.*, *8*(1), 54-63.

Kamei, F., Pinto, G., Wiese, I., Ribeiro, M., & Soares, S. (2021). *What Evidence We Would Miss If We Do Not Use Grey Literature?* Proceedings of the 15th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), Bari, Italy. https://doi.org/10.1145/3475716.3475777

Krivý, M. (2023). Digital ecosystem: The journey of a metaphor. *Digital Geography and Society*, *5*, 100057. https://doi.org/https://doi.org/10.1016/j.diggeo.2023.100057

Kvadsheim, H. D., Øvergaard, B. N., & Barstad, L. (2023). *IKT-sikkerhetstilstanden i kraftforsyningen 2023* Oslo, Norway: Norges vassdrags- og energidirektorat

LeSaint, J., Reed, M., & Popick, P. (2015). System security engineering vulnerability assessments for mission-critical systems and functions. 2015 Annual IEEE Systems Conference (SysCon) Proceedings,

Magutshwa, S. (2022). Rethinking the improvisation of digital health technology: A niche construction perspective. *Australasian Journal of Disaster and Trauma Studies*, *26*(Special Issue on Information Systems).

Magutshwa, S., & Radianti, J. (2021). A Qualitative Risk Identification Framework for Cyber-Physical-Social Systems. Proceedings of the 18th Information Systems for Crisis Response and Management Conference Blacksburg, VA, USA

Magutshwa, S., & Radianti, J. (2022). Is this Digital Resilience? Insights from Adaptation and Exaptation of a Cyber-Physical-Social System.

Mars, M. M., Bronstein, J. L., & Lusch, R. F. (2012). The value of a metaphor: Organizations and ecosystems. *Organizational Dynamics*, *41*(4), 271-280.

Mattos, D. I., Dakkak, A., Bosch, J., & Olsson, H. H. (2020). Experimentation for business-to-business mission-critical systems: A case study. Proceedings of the International Conference on Software and System Processes,

Meld.St.5. (2023). *En motstandsdyktig helseberedskap Fra pandemi til krig i Europa. Meld. St. 5 (2023–2024), Melding til Stortinget*. Helse-og-omsorgsdepatementet

Microsoft. (2023). Mission-critical workloads. *Azure*.

Motvind-Norge. (2019). *Energipolitikk på naturens premisser – til beste for klima, naturmangfold, mennesker og næringsliv*.

Mukherjee, M., Chatterjee, R., Khanna, B. K., Dhillon, P. P. S., Kumar, A., Bajwa, S., Prakash, A., & Shaw, R. (2020). Ecosystem-centric business continuity planning (eco-centric BCP): A post COVID19 new normal. *Progress in Disaster Science*, *7*, 100117. https://doi.org/https://doi.org/10.1016/j.pdisas.2020.100117

NORMEN. (2018). *Tiltak for å hindre ondsinnet programvare (faktaark 19)* Norway: Directorate for e-Health

NORMEN. (2021). *Nødprosedyrer ved bortfall av IKT (faktaark 11) , versjon 3.0*. Norway: Direktoratet for e-helse

NORMEN. (2022a). *Norm for informasjonssikkerhet og personvern i helse- og omsorgssektoren Versjon 6.1* Norway: Direktoratet for e-Helse

NORMEN. (2022b). *Veileder om risikostyring for informasjonssikkerhet og personvern* Norway: **Direktoratet for e-helse**

NORMEN. (2023). *Informasjonssikkerhet og personvern for leverandører til helse- og omsorgssektoren* Norway: Direktoratet for e-Helse

NOU. (2023a). *Mer av alt - raskere: energikommisjonens rapport*. Norway: Norges offenlige utredninger

NOU. (2023b). *Når det er alvor: Rustet for en usikker fremtid*. Norway: Norges offenlige utredninger

NSM. (2023). *Risiko 2023: Økt uforutsigbarhet krever høyere beredskap*. Norway: Nasjonal sikkerhetsmyndighet

Pavlikakis, G. E., & Tsihrintzis, V. A. (2000). Ecosystem Management: A Review of a New Concept and Methodology. *Water Resources Management*, *14*(4), 257-283. https://doi.org/10.1023/A:1008139011867

Pickett, S. T. A., & Cadenasso, M. L. (2002). The Ecosystem as a Multidimensional Concept: Meaning, Model, and Metaphor. *Ecosystems*, *5*(1), 1-10. https://doi.org/10.1007/s10021-001-0051-y

Reid, M. B. (2021). Business continuity plan. In *Encyclopedia of Security and Emergency Management* (pp. 52-57). Springer.

Rucks, A. C., Ginter, P. M., Duncan, W. J., & Lesinger, C. (2011). A continuity of operations planning template: Translating public policy into an effective plan. *Journal of Homeland Security and Emergency Management*, *8*(1), 000010220215477355177 5.

Schulz, F., & Rode, J. (2022). Public charging infrastructure and electric vehicles in Norway. *Energy Policy*, *160*, 112660.

Skjærseth, J. B., & Rosendal, K. (2023). Implementing the EU renewable energy directive in Norway: from Tailwind to Headwind. *Environmental Politics*, *32*(2), 316-337.

Thorne, R., Aguilar Lopez, F., Figenbaum, E., Fridstrøm, L., & Müller, D. B. (2021). Estimating stocks and flows of electric passenger vehicle batteries in the Norwegian fleet from 2011 to 2030. *Journal of Industrial Ecology*, *25*(6), 1529-1542.

Thorne, R. J., Hovi, I. B., Figenbaum, E., Pinchasik, D. R., Amundsen, A. H., & Hagman, R. (2021). Facilitating adoption of electric buses through policy: Learnings from a trial in Norway. *Energy Policy*, *155*, 112310.

Van der Meulen, N. (2013). Diginotar: Dissecting the first dutch digital disaster. *Journal of strategic security*, *6*(2), 46-58.

Veikart Helsenæringen. (2023). Norway: Helse- og omsorgsdepartementet

Veisten, K., Wangsness, P. B., Farstad, E., & Ydersbond, I. M. (2024). Will people prefer future travel with battery-powered airplanes? *Transportation Research Part D: Transport and Environment*, *126*, 104013.

Weger, K., Matsuyama, L., Zimmermann, R., Mesmer, B., Van Bossuyt, D., Semmens, R., & Eaton, C. (2023). Insight into user acceptance and adoption of autonomous systems in mission critical environments. *International Journal of Human–Computer Interaction*, *39*(7), 1423-1437.

Yip, J., Ernst, C., & Campbell, M. (2009). Boundary spanning leadership: Mission critical perspectives from the executive suite. *Center for Creative Leadership Organizational Leadership White Paper*.

Zuo, Y. (2013). Modeling service migration and relocation in mission-critical systems. Critical Infrastructure Protection VII: 7th IFIP WG 11.10 International Conference, ICCIP 2013, Washington, DC, USA, March 18-20, 2013, Revised Selected Papers 7,

Bridging Gaps, Building Futures: How DIREKTION Innovates for Global Disaster Resilience

Juliane Schlierkamp
Fraunhofer INT
juliane.schlierkamp@int.fraunhofer.de

Eileen Murphy
Trilateral Research
Eileen.murphy@trilateralresearch.com

Louise Vos
TNO
louise.vos@tno.nl

ABSTRACT

The DIREKTION project, funded under Horizon Europe, aims to revolutionize disaster resilience through advanced technology and collaboration. Addressing the challenges posed by climate change, infrastructure vulnerabilities, and dynamic population densities, it seeks a paradigm shift in emergency responses. By integrating innovative technologies with the practical needs of responders and policymakers, DIREKTION enhances cross-sectoral harmonization and fosters stakeholder communication. Early outputs include drafts of the DIREKTION Assessment and Screening Framework and D-SAT tool, establishing a foundation for future development. The project's presentation at ISCRAM invites valuable feedback and engagement, highlighting its commitment to improving disaster management practices and fostering a more resilient society.

Keywords

Disaster resilience, Capability-based approach, Stakeholder collaboration, Emergency response management.

INTRODUCTION

In response to the increasing complexity and frequency of disasters, the DIREKTION project was initiated in October 2023 and is scheduled for a 36-month duration. Funded under the Horizon Europe's cluster 3 security call HORIZON-CL3-2022-SSRI-01, the project aims to improve disaster resilience through a combination of technological innovation and collaboration among multiple stakeholders. The need for DIREKTION arises from multiple challenges, including climate change, aging infrastructure, geopolitical instabilities, and the complexities of managing dense, dynamic populations. These challenges necessitate a new approach to emergency responses, which DIREKTION seeks to address by fostering better knowledge sharing and developing technologies that address the needs of emergency responders and policymakers.

The project focuses on integrating innovative technologies into disaster management practices through thorough assessments of existing solutions, identifying gaps, and prioritizing future research and innovation efforts. It utilizes outputs from previous EU Horizon projects and creates a network of industry stakeholders, small and medium-sized enterprises (SMEs), research organizations, and practitioners to build a more resilient and prepared society.

This paper presents the DIREKTION project, outlining its goals, methodology, and expected impact on disaster resilience and crisis management. We invite the research and practitioner communities to join a collaborative dialogue to improve our collective ability to anticipate, respond to, and recover from disasters.

BACKGROUND

Facing escalating challenges like climate change, technological disruptions, and societal shifts, there is an urgent need to enhance civil security and crisis management capabilities. The European Union has launched the Horizon Europe call HORIZON-CL3-2022-SSRI-01 under its flagship research and innovation program, aimed at advancing research and innovation in this field (European Commission). One of the projects funded under this initiative is DIREKTION, designed to bolster disaster resilience and management using innovative technologies and methodologies. Horizon Europe's Cluster 3, titled "Civil Security for Society," aims to improve security through solutions for combating crime and terrorism, managing borders, enhancing cybersecurity, and building disaster-resilient societies (European Commission). The DIREKTION project supports these goals by adopting an integrative approach to crisis management that emphasizes advanced technology and cross-sectoral collaboration.

The increasing impacts of climate change, evidenced by more frequent and severe weather events (Kron et al., 2019, p. 82), alongside new technologies, including those developed to cope with climate change, such as electric mobility (Liu et al., 2023, p. 9), that present new firefighting challenges, underscore the urgency of these initiatives. These developments complicate traditional disaster response mechanisms and underscore the need for a flexible, comprehensive approach to civil security. The mission of the DIREKTION project aligns with Horizon Europe's broader goal of creating innovative, inclusive, and sustainable solutions to contemporary challenges. Strengthened Security Research and Innovation (SSRI), a key focus of the program, advocates for transcending sector-specific biases and dismantling existing silos that impede the spread of common security solutions (European Commission).

DIREKTION's integrative approach to crisis management emphasizes advanced technology and cross-sectoral collaboration. By connecting a "network of networks" that includes established fire and rescue organizations, academics, research entities, industry stakeholders, and policymakers, DIREKTION facilitates a comprehensive strategy to address capability needs and gaps, technology screening and sourcing, assessment of technology and project results, co-development of technologies, research programming, and practical standardization aligned with practitioner needs. Additionally, acknowledging the critical role of human factors in crisis management, DIREKTION focuses on designing user-centric solutions tailored to the requirements and capabilities of first responders and affected communities, aiming to enhance both technological and social capacities for effective crisis response.

This collaborative model, supported by Horizon Europe’s framework, underlines the significance of international cooperation and knowledge sharing in addressing global security challenges. DIREKTION exemplifies the program’s vision for a unified European research area focused on societal safety and well-being. Reflecting on DIREKTION's alignment with the objectives of Horizon Europe’s Cluster 3 demonstrates that the project is not just a response to a funding opportunity but a strategic effort to push the boundaries of civil security research and innovation, embodying the EU’s dedication to using scientific research for societal benefits and setting new standards in disaster resilience and crisis management.

METHODOLOGY

In this paper, we detail the methodology adopted by the DIREKTION project, aimed at addressing the fragmented nature of the research projects and the small-scale, highly fragmented security market. Recognizing the limitations of existing taxonomies based solely on technological solutions, DIREKTION employs a capability-based approach to enhance cross-sectoral and cross-departmental harmonization, improve access to research results and market solutions, and foster communication among interdisciplinary stakeholders in civil security. This methodology leverages the results of prior research (e.g. ACRIMAS[1], ResiStand[2], FIRE-IN[3], MEDEA[4]) and integrates the 13 functional areas identified in the European Commission's study of the EU security market (Deloitte LLP et al., 2022), facilitating a comprehensive mapping of capability needs and solutions.

The project is structured into seven interdependent work packages (WPs) (see Figure 1), each with specific objectives.

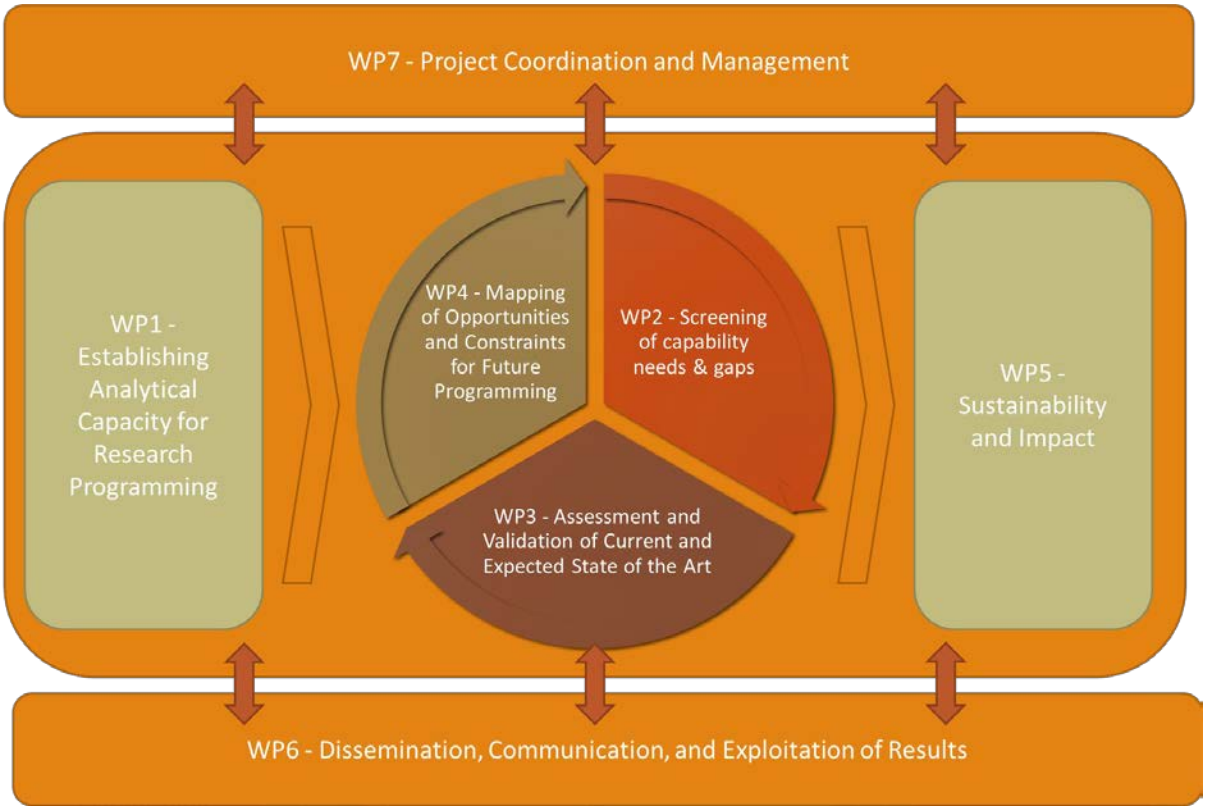


Figure 1. work package structure of DIREKTION

In *WP1 - Establishing Analytical Capacity for Research Programming*, the cornerstone of DIREKTION's methodology, a common, sustainable framework that underpins the entire project is being developed. This work package draws upon a wide range of sources, including previous research findings and the taxonomy of the EU Security Market study, to establish a comprehensive categorization system. This system is designed to facilitate the identification and classification of capability needs and solutions, thereby addressing one of the critical challenges: the fragmentation of the security market. By developing a common language and framework for capability needs, WP1 lays the foundational groundwork for the project's success, enabling more effective communication and collaboration among stakeholders.

The WPs 2 to 4 built an iterative process of 3 cycles over the project duration, starting with WP2. *WP2 - Screening and Mapping of Capability Needs & Gaps* focuses on the meticulous identification of current and future capability needs and gaps within the security domain. This involves a thorough analysis of operational requirements, technological advancements, and emerging threats to ensure that the project's efforts are squarely targeted at areas of critical need. WP2's objective is to systematically map out these needs and gaps, providing a clear direction for research and innovation activities. This mapping process is essential for prioritizing research efforts and ensuring

[1] https://cordis.europa.eu/project/id/261669
[2] https://cordis.europa.eu/project/id/700389
[3] https://cordis.europa.eu/project/id/740575
[4] https://cordis.europa.eu/project/id/787111

that the project's outcomes are both relevant and impactful in addressing real-world security challenges.

Based on the WP2 outcome *WP3 - Assessment and Validation of Current and Expected State of the Art* is pivotal in evaluating the operational capacity of existing and forthcoming solutions against identified capability needs. This work package employs a rigorous validation process to assess the alignment, efficacy, and potential of solutions to meet the articulated needs. By involving practitioners and stakeholders in this assessment, WP3 ensures that the solutions are not only technologically advanced but also practically viable and tailored to operational contexts. This alignment between solutions and needs is crucial for the successful operationalization and adoption of new technologies in the security sector.

WP4 - Mapping of Opportunities and Constraints for Future Programming undertakes a comprehensive analysis of the factors that influence the successful implementation and uptake of new solutions under consideration of the WP2 and WP3 findings. This includes examining market drivers, barriers to standardization, and the ethical and societal implications of security innovations. By identifying these opportunities and constraints, WP4 provides valuable insights that inform the strategic direction of future research programming. This work package plays a crucial role in ensuring that DIREKTION's efforts are not only technologically sound but also market-ready and socially responsible.

WP5 - Sustainability and Impact is dedicated to fostering innovation and enhancing multi-stakeholder cooperation within the DRS community. Through the establishment of networking activities and the introduction of awards for innovation, WP5 aims to build a vibrant ecosystem that supports the exchange of ideas, resources, and best practices. This work package is instrumental in promoting the uptake of new technologies and methodologies, ensuring that DIREKTION's outcomes have a lasting impact on the field of civil security.

WP6 - Dissemination, Communication, and Exploitation of Results seeks to amplify the reach and impact of DIREKTION's findings and innovations. By implementing a targeted communication strategy, this work package ensures that the project's results are effectively shared with a broad audience, including industry partners, SMEs, researchers, and practitioners. It should be noted here that the aim is an exchange and not a one-dimensional flow of information. WP6 plays a vital role in establishing a platform for the dissemination and exploitation of solutions, facilitating the transfer of knowledge and technology from the research domain to operational practice.

WP7: Project Coordination and Management ensures the seamless execution and governance of the DIREKTION project. Entrusting the steering role to practitioners from the consortium, this work package embodies the project's commitment to a capability-driven approach. By involving practitioners in the leadership and decision-making processes, WP7 guarantees that the project's developments are closely aligned with the needs and realities of the DRS community, fostering innovation that is both relevant and impactful. Our approach emphasizes a capability-driven strategy for developing new solutions, involving practitioners early in the research process to ensure that future innovations meet operational requirements and market expectations.

By incorporating a mix of foresight capabilities, operational assessments, and strategic cooperation, DIREKTION aims to bridge the gap between demand and supply, facilitating the development of solutions that are both innovative and practically applicable. This methodology not only aims to overcome market fragmentation but also establishes a sustainable framework for enhancing civil security research and innovation, ensuring that future projects are aligned with the real-world needs of practitioners and stakeholders.

Our approach emphasizes a capability-driven strategy for developing new solutions, involving practitioners early in the research process to ensure that future innovations meet operational requirements and market expectations. By incorporating a mix of foresight capabilities, operational assessments, and strategic cooperation, DIREKTION aims to bridge the gap between demand and supply, facilitating the development of solutions that are both innovative and practically applicable. This methodology not only aims to overcome market fragmentation but also establishes a sustainable framework for enhancing civil security research and innovation, ensuring that future projects are aligned with the real-world needs of practitioners and stakeholders.

PRELIMINARY RESULTS AND STATE OF THE PROJECT

The DIREKTION project, in its initial stages, has made considerable progress in developing its foundational elements, namely the DIREKTION Assessment and Screening Framework and the associated DIREKTION Solution Assessment Tool (D-SAT) drafts and its supporting user guide. These preliminary outputs are critical in shaping the project's approach to identifying capability needs and gaps within the domain of Disaster Resilient Societies (DRS), as well as in screening and assessing existing and state-of-the-art technologies. Although these outputs are currently in draft form, they mark a significant step towards establishing a structured methodology for enhancing disaster resilience through innovative technology uptake.

DIREKTION Assessment and Screening Framework (DASF)

The DIREKTION Assessment and Screening Framework (DASF) has been developed within DIREKTION's work package 'Establishing Analytical Capacity for Research Programming' (WP1). It resulted in the initial version of a robust framework of methods for identifying capability needs and addressing gaps, as well as implementing screening and assessment processes, to be used for programming future research.

To create the initial DASF, existing methods from both EU projects and external sources are screened beforehand. Methods that align with the needs of the DIREKTION project, either partly or fully, are integrated into the DASF. One of the challenges in developing the DASF was dealing with the complexity of how disaster management is organised, as EU Member States have their own organisational crisis and disaster management structures, culture, traditions and administrative arrangements which are not expected to be changed easily. This has caused a lack of common standardized language in the disaster resilience domain in the EU. Therefore, to identify and describe capability needs and gaps of the EU and all its Member States, a harmonized and well-balanced approach is needed which takes into account the differences in the way they have organized disaster management and the need for standardized terminology. While reviewing existing methods, it became evident that the ResiStand[5] project (H2020) developed a generic framework for disaster management tasks (Stolk et al., 2017). This framework covers all phases of the disaster management cycle – preparedness, response, recovery, mitigation – with the aim of establishing standardized terminology and practices. This framework from ResiStand helped shape the DASF. For practical reasons scoping within the disaster management cycle was needed. The current DASF mainly focuses on the preparedness and response phase instead of elaborating on the entire disaster management cycle; as these are typically the phases in which first response organizations are involved in. Another important role has the FIRE-IN methodology. The FIRE-IN project aimed to enhance the capability development process in the Fire & Rescue sector across Europe. The FIRE-IN method, at its core, consists of an iterative process of identifying capability gaps by workshops and assessing solutions. (Sakkas et al., 2023)

The DASF consists of a set of four methods that include ten steps in total. In Figure 3 these steps are shown as a linear set of subsequent steps. The complete set of steps is meant to be executed regularly, to get an updated insight in the most actual capability needs, capability gaps, potential solutions and to regularly update the roadmap for EU-Research opportunities. As such, the DASF provides the basis for a sustainable process for research programming.

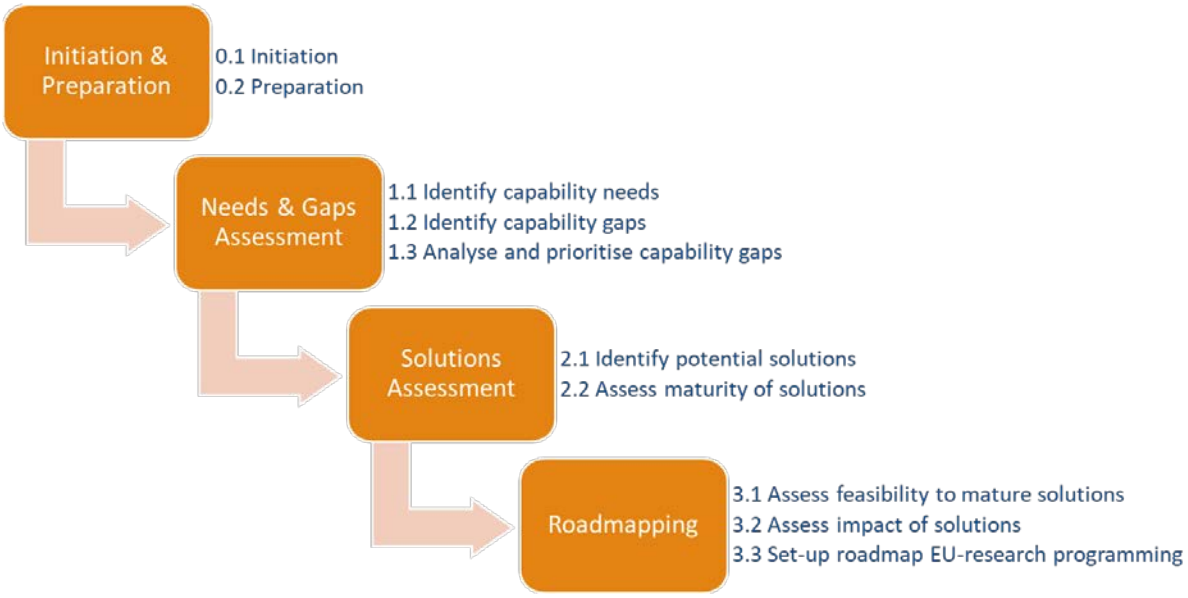


Figure 2: Overview of the methods and steps in the DASF

The initial phase, "Initiate and Prepare," is divided into two sub-steps:

- Step 0.1 – Initiation: This step involves preliminary activities that define the framework for the subsequent assessment cycle. Key actions include assigning roles and responsibilities to participants, determining the scope of the assessment (e.g., general disaster response or focusing on specific threat mitigation), and overall project planning.

[5] https://cordis.europa.eu/project/id/700389

- Step 0.2 – Preparation: Following initiation, this step involves preparatory actions to ensure readiness for further assessment and screening in later steps. Activities include identifying, selecting, and informing stakeholders, developing necessary documentation, and preparing supporting tools for steps 2 (Needs and Gaps Assessment) and 3 (Solution Assessment).

The Needs and Gaps Assessment comprises:

- Step 1.1 – Identify capability needs: Using risk and trend analyses, historical accident data, scenarios, and consultations with end-users and experts, this step gathers capability needs for disaster resilience and security (DRS) activities within the EU, for the present and future.
- Step 1.2 – Identify capability gaps: This involves comparing existing disaster management capabilities with the identified needs to highlight capability gaps.
- Step 1.3 – Analyse and prioritise capability gaps: The identified gaps are analyzed to assess their impact on current and future DRS capabilities, determining which gaps should be prioritized for resolution.

Solution Assessment includes:

- Step 2.1 – Identify potential solutions for bridging the gaps: This step involves gathering information on potential solutions, including mature solutions available on the market, as well as prototypes and promising research activities.
- Step 2.2 – Assess maturity of potential solutions: The maturity of these solutions is evaluated based on technical, societal acceptance, integration, and commercialization aspects.

The Roadmapping phase is structured into three sub-steps:

- Step 3.1 – Assess feasibility to mature potential solutions: For solutions still in development, this step assesses the feasibility of developing these into fully mature solutions, including potential challenges toward commercial availability.
- Step 3.2 – Assess impact of potential solutions: This step evaluates the potential impact of solutions on bridging one or more capability gaps.
- Step 3.3 – Set-up roadmap for EU-Research Programming: Aims to propose a research programming roadmap for the short (3 years), medium (5-10 years), and long terms (15+ years), integrating findings from the previous steps to plan future DRS research in Horizon Europe.

DIREKTION Solution Assessment Tool (DSAT)

DSAT is designed to support emergency management organisations in systematically comparing and assessing multiple solutions in the context of identified and prioritised capability gaps. The tool builds on the core principles of maturity models. Maturity models have been successfully applied across a range of fields to assess capability. A maturity model is “essentially a classification scheme that places patterns in developing organizational capabilities under a certain capability stage, assuming linear progression from an existing mature state to a higher maturity level” (United Nations Development Programme [UNDP], 2022, p. 10)

D-SAT is an excel based tool that is designed to be easily accessible to multiple users. It involves four components that can be flexibly combined to suit assessment needs. The different components are targeted at different stakeholder groups, with the overarching aim of supporting collaboration and dialogue between supply and demand side actors, successful solution development, and market entry and uptake. D-SAT focuses on two key stakeholder groups. Solution providers (see Figure 2, in orange) and emergency management organisations (Figure 2, in blue) and their interaction in the context of innovation (orange and blue).

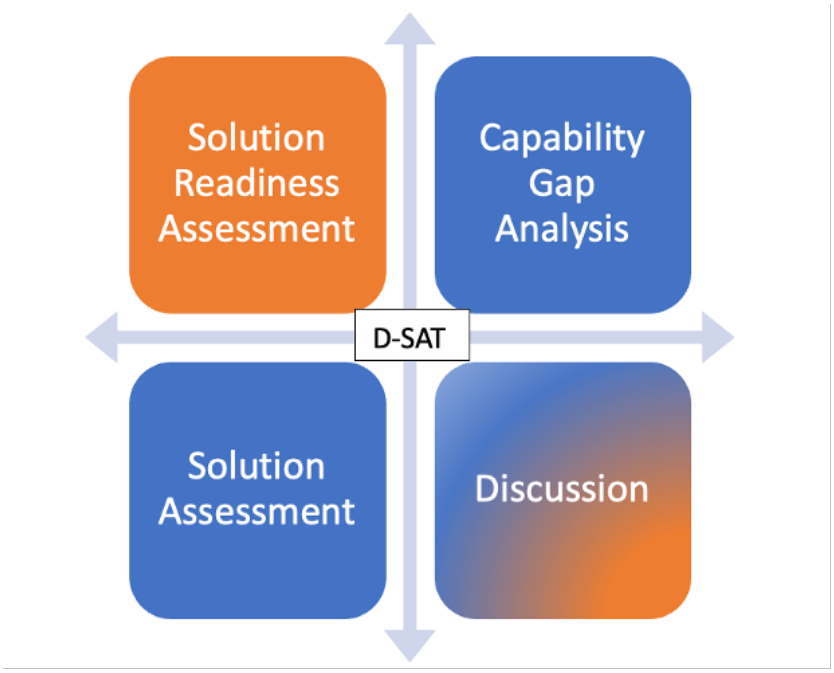


Figure 3: D-SAT Components by Stakeholder Type

The 'Solution Readiness Assessment' component is completed by the solution provider and provides information on the readiness of the solution on a variety of scales: technology; societal; manufacturing; integration; commercialisation; legal, privacy, and ethical; and security. This component is intended to provide a structured space to capture information on the readiness of the solution and builds on the ongoing work of the EU-funded project MultiRATE[6].

In the 'Capability Gap Analysis' component the assessment owner will score identified capability gaps using a maturity scale (from 0-5). The maturity of capability gaps is analysed by capability elements: infrastructure; interoperability; technology; human; organisation; regulatory., building on the results of the EU-funded project MEDEA. Results are then assigned a complexity domain according to the Cynefin framework (Snowden & Rancati, 2021) and matched with potential solution functionalities drawing on the results of the EU Security Market study. (Deloitte LLP et al., 2022). This analysis of capability gaps will provide a baseline assessment and allow users to determine the level of response required when choosing and investing in a solution.

In the 'Solution Assessment' component the assessment owner score solutions (from 1-5) against two sets of criteria. The first assesses the likelihood that a solution(s) will improve the maturity of your capability elements. The second assesses the importance of several high-level priorities for the adoption and implementation of the solution being assessed. These are grouped according to three intersecting topics: 1) legal, societal and ethical; 2) user needs; and 3) market. These groupings are flexible and often overlap with one another. The level of importance of these factors varies depending on the kind of solution being assessed. For example, the relationship between these factors and solutions based on AI technologies is rapidly evolving. At the same time changes in supply chains and the shift to thinking in terms of the environmental costs of manufacturing and production are changing how we assess the value of current solutions. Thus, the importance of these factors is linked to the context of the specific solution assessment and its expected use case.

In the 'Discussion' component a range of questions are presented to promote discussion about the factors that enable and hinder the adoption and implementation of solutions. These questions are grouped according to three broad topics: 1) Solution Adoption and Implementation; 2) Solution Impact; and 3) Innovation collaboration, with a focus on EU research. These questions have been adapted from previous DRS research (e.g. ResiStand) and designed to promote the process of reflecting on the value of engaging in innovation collaboration and the adoption of innovation outcomes and results. And to support the identification of potential barriers and enables for the uptake of research and innovation results.

If the solution is on or close to market, the assessment results can help to guide and refine the procurement process. The assessment results should allow emergency management organizations to see if the solution will meet their

[6] https://cordis.europa.eu/project/id/101073929

capability gap needs and provide guidance on how best to use resources. If the solution is still under development, the results can be used to guide collaboration and engagement with providers on what is needed next, either for the solutions assessed or future needs identification. A comparison of the readiness levels assigned to the solution by the solution provider and the results of the assessment will help to see how well supply and demand perspectives are aligned. Finally, repeating the assessment process will allow emergency management organizations to check if the solution has improved their maturity score over time, or if further investment is required. Repeating the capability gap analysis, even without a solution to assess, will also allow emergency management organizations to track how their capability gap needs have evolved over time.

Extant literature has highlighted the role innovation solutions play in the disaster management, especially in the face of an evolving risk landscape (Izumi, et al. 2019; Rahman and Fang, 2019). Research has also demonstrated the bi-fold challenges related to 1) bringing innovative solutions to market (Butler, 2008) and 2) achieving meaningful uptake and adoption of solutions. This is further complicated by a lack of frameworks dedicated to the assessment of innovation potential of solutions (Klessova et al, 2022). Add to this, the fact that responder organisations are traditionally slow to adapt their existing practices and techniques, there is a well-established need for supportive tools to improve demand and supply interaction in the context of innovation assessment (Webster and Gardener, 2019; Nepelski and Van Roy, 2021).

DSAT aims to meet this need by providing a flexible tool to develop and strengthen collaborative solution assessment and better match supply and demand side needs (see Krikigianni et al, 2022). Existing efforts to do so have focused either on a specific solution type, primarily technology, or domain of application, e.g. flooding, wildfires. However, DSAT has been designed to be applied to a range of solution types, including methods and standards. In addition, it adopts an all-hazards approach, providing further flexibility.

The application of the tool will be tested in the context of the DIREKTION project to address a range of solution types across three stages of innovation readiness: 1) close to or on the market solution, 2) under development solutions, and 3) early-stage solutions. The assessment of these solutions will be matched to the identification and prioritization of current and future capability gaps. DSAT seeks to go beyond state-of-the-art approaches through the integration of existing best practices to develop a generalizable framework that will allow for a systemic assessment of solutions and provide a baseline for comparison of solutions across assessing organisations and actors.

As we prepare for the presentation at the International Conference on Information Systems for Crisis Response and Management (ISCRAM), these early results highlight the project's potential impact and its commitment to addressing the complex needs of disaster management. Presenting DIREKTION at this stage offers a dual benefit: it allows for the integration of valuable feedback from a diverse audience of experts, practitioners, and researchers, and it opens avenues for active engagement and collaboration within the project's growing network. This early dialogue is essential for refining the project's outputs and ensuring their relevance and applicability to the challenges faced by the DRS community.

By sharing these preliminary results, we invite constructive feedback and collaboration, laying a foundation for the iterative development of the DIREKTION project. This approach ensures that the project remains responsive to the evolving needs of disaster management and maximizes its contribution to the field.

CONCLUSION

The DIREKTION project marks a significant advance in the field of disaster resilience and crisis management. As global societies face an increasing frequency and complexity of disasters, driven by factors such as climate change, technological advancements, and geopolitical shifts, there is a critical need for innovative approaches that enhance the capabilities of communities and emergency responders. Funded under the Horizon Europe framework, the DIREKTION project, part of the "Civil Security for Society" cluster, aims to fill current technological and operational gaps in disaster response while promoting a culture of preparedness that extends beyond traditional sectoral and national boundaries.

At the core of the project's methodology is a capability-based approach, which shifts the focus from reactive to proactive and predictive disaster management strategies. This approach involves evaluating existing solutions, pinpointing capability gaps, and focusing research and innovation efforts to develop technologies that meet the practical needs of frontline responders and policymakers. Such alignment is essential to ensure that technological advancements lead to real improvements in operational capacity during disaster responses.

DIREKTION is structured around seven work packages that collectively aim to achieve its ambitious objectives. These packages range from developing analytical capacities for research programming to enhancing sustainability and impact through multi-stakeholder cooperation. The project places a strong emphasis on dissemination, communication, and the exploitation of results to ensure the broad application and maximal impact of the

innovations developed.

Initial outcomes include the development of the DASF Framework and the D-SAT tool, which are major steps forward in structuring disaster management practices. The DASF Framework provides a systematic approach to research programming, while the D-SAT tool enables emergency management organizations to evaluate technological solutions against specific capability gaps. These achievements illustrate early successes in linking technological innovation with operational needs.

In conclusion, the DIREKTION project contributes significantly to enhancing disaster resilience. Through its methodological rigor, collaborative structure, and commitment to incorporating stakeholder feedback, DIREKTION not only pushes forward the state-of-the-art in crisis management but also builds foundations for a more resilient and capable society. As the project progresses, its focus on stakeholder engagement, capability-driven research, and feedback integration will be key to its ongoing success and relevance in an ever-changing world.

REFERENCES

Butler, D. (2008). Translational Research: Crossing the Valley of Death. Nature, 453, 840-842.

Deloitte LLP; Ecorys; Europäische Kommission. (2022). *Eu security market study: Final report : 31 May 2022*. Publications Office of the European Union. https://op.europa.eu/o/opportal-service/download-handler?identifier=db2efbc8-070a-11ed-acce-01aa75ed71a1&format=pdf&language=en&productionSystem=cellar&part= https://doi.org/10.2837/19472

European Commission. *Horizon Europe - Cluster 3: "Civil security for society"*. https://rea.ec.europa.eu/funding-and-grants/horizon-europe-cluster-3-civil-security-society_en

European Commission. *Strengthened Security Research and Innovation (SSRI)*. https://rea.ec.europa.eu/funding-and-grants/horizon-europe-cluster-3-civil-security-society/strengthened-security-research-and-innovation-ssri_en

European Commission. *Strengthened Security Research and Innovation (SSRI)*. https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/horizon-cl3-2022-ssri-01-01v

Izumi,T., Shaw, R., Djalante, R., Ishiwatari, M., & Komino, T. (2019). Disaster risk reduction and innovations. Progress in Disaster Science. https://doi.org/10.1016/j.pdisas.2019.100033Get rights and content.

Klessova, S., Engell, S. & Thomas, C. (2022). Assessment of the advance of market-stream innovations and of the performance of research and innovation projects. Technovation. https://doi.org/10.1016/j.technovation.2022.102495

Krikigianni, E., Latsa, E., Gayialis, S., Panayiotou, N., Kostovasili, M., Kanellopoulos, I. & Amditis, A. (2022). Matchmaking the Emerging Demand and Supply Need in the Maritime Supply Chain Domain: A System Design Framework. Sustainability. https://doi.org/10.3390/ su142114622

Kron, W., Löw, P., & Kundzewicz, Z. W. (2019). Changes in risk of extreme weather events in Europe. *Environmental Science & Policy*, *100*, 74–83. https://doi.org/10.1016/j.envsci.2019.06.007

Liu, J., Xu, N., Shi, Y., Barnett, T., & Jones, S. (2023). Are first responders prepared for electric vehicle fires? A national survey. *Accident; Analysis and Prevention*, *179*. https://doi.org/10.1016/j.aap.2022.106903

Nepleski, D., & Van Roy, V. (2021). Innovation and innovator assessment in R&I ecosystems: the case of the EU Framework Programme. The Journal of Technology Transfer. https://doi.org/10.1007/s10961-020-09814-5

Rahman, A. & Fang, C. (2019) Appraisal of gaps and challenges in Sendai Framework for Disaster Risk Reduction priority 1 through a framework of science, technology and innovation. Progress in Disaster Science. https://doi.org/10.1016/j.pdisas.2019.100006Get rights and content

Sakkas, G., Tsaloukidis, I., Kazantzidou-Firtinidou, D., Schneider, I., Kouskouna, V., Hybbeneth, N., Berchtold, C., Schlierkamp, J., Miralles, M., Lahaye, S., & Bour, M. (2023). The FIRE-IN project: Tsunami-risk related practitioner challenges and 3rd cycle overall results. *Open Research Europe*, *3*, 5. https://doi.org/10.12688/openreseurope.15249.2

Snowden, D., & Rancati, A. (2021). *Managing complexity (and chaos) in times of crisis: A field guide for decision makers inspired by the Cynefin framework*. *EUR: Vol. 30569*. Publications Office of the European Union.

Stolk, D., van der Lee, M., Petiet, P., Liedtke, C., Lindner, R., van Zetten, J., Karppinen, A., Anson, S., & Kroener, I. (2017). *Assessment framework for standardisation*. https://cordis.europa.eu/project/id/700389/results

United Nations Development Programme (UNDP). (2022). *Digital Disaster Risk Reduction Maturity Model (DDRRMM)*. https://www.undp.org/sites/g/files/zskgke326/files/2022-09/DDRRMM%20White%20Paper%20Version%204.0%20%28FINAL%29_1%20September%202022.pdf

Webster, A. & Gardener, J. (2019) Aligning technology and institutional readiness: the adoption of innovation. *Technology Analysis & Strategic Management.* https://doi.org/10.1080/09537325.2019.1601694

Pre-disaster Training for Firefighters Using Augmented Reality: Hologram Characteristics

Stella Polikarpus
Estonian Academy of Security Sciences
stella.polikarpus@sisekaitse.ee

Reet Kasepalu
Estonian Academy of Security Sciences
reet.kasepalu@sisekaitse.ee

Kärt Reitel
Estonian Academy of Security Sciences
kart.reitel@sisekaitse.ee

ABSTRACT

The utilization of Augmented Reality (AR) holograms in firefighters' training presents a significant opportunity for the acquisition of first responders' practical skills. Nevertheless, given that AR represents a nascent and evolving technological paradigm, its integration within the framework of the crisis management cycle has not yet been firmly established. We analysed three hologram characteristics: dimension, scalability, and interactivity and report on the results of using five holograms (one 2D and four 3D) in a training application FightARs. The holograms were tested among in-service firefighters and commanders with the purpose of improving their knowledge in pre-disaster traffic accident training involving hazardous materials. The results indicate that QR-code scanning to open non-interactive non-scalable holograms can be used effectively to integrate holograms into physical training environments. A full-scale holographic scene is useful for training situation awareness. We propose that 3D non-interactive scalable or non-scalable holograms should be used to exercise practical skills. The paper emphasizes the necessity of extensively testing and incorporating AR into the training of first responders prior to its application in the response phase of crisis management.

Keywords

Augmented Reality (AR), Hologram types, FightARs, Holographic scene, firefighters' training

INTRODUCTION

Historically, a variety of methods have been employed to enhance the training process through the visual presentation of information. Many technologies are used to create and display 3D visual content (Ahmad et al., 2021) and one of the most important 3D display technologies is an Augmented Reality (AR) device which allows to augment a three-dimensional virtual element into the real world (Rauschnabel et al., 2015) such as HoloLens 2. The utilization of AR in firefighters' training presents a significant opportunity for the acquisition of practical skills. Looking ahead, it is anticipated that augmented reality (AR) devices will become more cost-effective, reliable, and proficient in data management, while Wi-Fi technology is expected to achieve greater speeds. Artificial intelligence (AI) could be effectively employed for the rapid analysis of multimodal data derived from algorithms and sensors. Nevertheless, the integration of augmented reality (AR) technology within the crisis response phase is currently ready solely for training purposes and remains a target yet to be achieved in real-world applications. In the project RESCUER, a smoke-removal algorithm, an object detection model and AR HoloLens 2 application were used for visualisation purposes (Gkika et al., 2023). In the FASTER project, the same device

was used to improve situation awareness by marking the points of interest (Christaki et al., 2022). In this project, the experiences created with HoloLens 2 were called "a real Mixed-Reality experience" (ibid p. 758). The RESCUER project was designed to enhance the capability to locate individuals in smoky environments, while the FASTER project focused on identifying points of interest on maps and employing HoloLens in the real world. These aforementioned applications developed for first responders were meant to be used in the response phase of crises. Notwithstanding, the learning experience was not the main focus of these studies. The authors simply noted that the HoloLens applications required some initial practice by the majority of participants, due to the participants lack of prior experience with the developed technology (Christaki et al., 2022).

To successfully integrate any new technology into educational settings, three critical dimensions must be considered: the technology, the pedagogical approach, and the learning content (Kurt, 2019). Firefighters use simulation-based training extensively (Polikarpus et al., 2021), and the trainer's role of choosing a suitable pedagogical approach for virtual simulation-based training has been highlighted (Keskitalo, 2022; Polikarpus, Luik, et al., 2023). If trainers are unable to operate the technology and have not acquired the necessary vocabulary for instructing trainees, they might experience difficulties in conducting situated learning (Herrington & Oliver, 1995). Furthermore, "hologram technology differs from augmented reality in several aspects, as the way in which the three-dimensional images are created and the properties of the images itself (Elmahal et al., 2020)". Even if the use of holography in education has a positive trend in last five years, the studies are limited to small age groups, subjects, and the technological support needed to achieve learning outcomes are still unclear (Yoo et al., 2022).

Therefore, it is important for the ISCRAM community and the trainers of first responders to further study using new technologies during training so that they could be more effectively used to prepare for future crises. Preparedness activities, primarily training and public education within pre-disaster hazard management, are crucial for enhancing survival during disasters when effectively implemented (Bullock et al., 2013). Processes like preparing, teaching, simulating, and training are essential for an efficient transfer from the post-disaster phase to the pre-disaster phase in the crises management cycle (Aligne & Mattioli, 2011).

This paper is organized as follows: initially, we present the related work on using AR in education and describe hologram characteristics. After this, we present the holograms used in the FightARs project in four case studies. the ensuing section, we explore the integration of HoloLens 2 hologram types with different phases of crisis management, suggesting the appropriate hologram for each crisis management phase.

RELATED WORK

Mitigation, preparedness, response and recovery are seen as the four classical phases of disaster management (Bullock et al., 2013). To emphasize the significance of context, the initial two phases have been combined into a single phase, resulting in three phases (see Figure 1): pre-disaster, response, and post-disaster (Aligne & Mattioli, 2011). To the best of our knowledge, the potential of firefighters using AR has so far been focusing on the response phase of crises (Christaki et al., 2022; Gkika et al., 2023), as this is where it has the biggest potential to save lives by improving the situation awareness of first responders. Nevertheless, the training needs of the users of AR were not studied.

In education, three-dimensional (3D) learning environments are used as it gives a possibility to view a given problem from different perspectives and include virtual activities that are difficult to practice safely in real life (Reisoğlu et al., 2017). Virtual reality (VR) is the use of computer-based simulation enabling a user to interact with a 3D synthetic environment which often simulates a real world or imaginary scenario (He & Zhu, 2022). VR offers possibilities of making trainings both safer as well as repeatable when compared to other methods that have proven to be expensive or pose considerable risks to firefighters (Heldal & Wijkmark, 2017; Polikarpus et al., 2019). Virtual simulation-based training has been shown to be an effective way to train dynamic decision-making across all levels of crisis management (Lamb et al., 2021) and immersive VR based simulation for training firefighting skills are being developed for simulated fire extinguishing (Wijkmark et al., 2021, 2022). However, a significant limitation of virtual simulation-based training, whether utilizing virtual reality or computer-simulated videos, is that trainees are immersed in a virtual environment rather than interacting with the real world. Consequently, trainees are unable to practice their skills in virtual reality with the same level of practical application as they would in real-life scenarios.

AR mainly consists of the real world and is supplemented by virtual images generated by a computer, (Ma & Choi, 2007) and it has not gained widespread adoption in firefighter training programs (Hofmann & Polikarpus, 2022). A *hologram* is a special type of 3D object which can be seen with the eye and there are various types of hologram technologies (Elmahal et al., 2020). Yoo et al. (2020) propose that contemporary applications of

holography in educational settings could be categorized into three distinct types: (a) pseudo-holograms, (b) analog holograms, and (c) digital holograms and they attributed all AR and VR device-assisted holograms into digital holograms. Işik, (2014) introduces the concept of "dimensional holograms," defining a 1D hologram as a linear recording along one of the x, y, or z axes. Conversely, a 3D hologram is described as a luminous, transparent, three-dimensional image that is visible from all viewpoints and can be physically circumnavigated. Consequently, all studied holograms in this paper are 2D or 3D digital holograms produced with augmented reality glasses Hololens 2. Furthermore, holograms generated by the HoloLens 2 invariably contain data along the x, y, and z axes. As a result, merely distinguishing between 2D and 3D holograms does not provide software developers with sufficient information to accurately classify the type of hologram.

Holographic-type communication (HTC) will facilitate the transmission and interaction with holographic data across a network from remote locations (Clemm et al., 2020). HoloLens 2 already enables HTC when the “Dynamics 365 Remote Assist” app is used (Microsoft, 2024). For HTC, it is crucial to understand the specific characteristics that the transmitted holograms should possess. During the study of the FightARs application, the digital twin served as an object that allowed firefighters to interact with it virtually (The Fightars Partnership, 2021), whereas the scientific definition of a digital twin is that it “is a virtual representation of a physical object or process capable of collecting information from the real environment to represent, validate and simulate the physical twin’s present and future behaviour (Botín-Sanabria et al., 2022, p. 1)”.

Prior research did not give any suggestions on which hologram characteristics help to train firefighters knowledge and skills using HoloLens 2. Consequently, end-users of HoloLens 2, such as firefighters and trainers of first responders, lack the ability to specify to developers the types of holograms required for future training sessions, assessments, or the response phase to improve knowledge and skills in crisis management.

HOLOGRAM CHARACTERISTICS AND DATA COLLECTION

The FightARs project aimed to introduce a new AR device HoloLens 2 to the training of firefighters and rescue incident commanders (FightARs, 2022). We aim to answer the following research question: “**what type of holograms supports firefighters' knowledge and skills' training in which crisis management phase?** We categorized the holograms used in the FightARs' application based on three characteristics: *dimension* (2D or 3D), *scalability and interactivity* (see Figure 1 and Table 1). In our case study we linked HoloLens 2 hologram characteristics and firefighters' preparedness to use AR in the crisis management cycle taking account the current technological level in firefighters' training in Estonia.

For our case study, we used the results from four different empirical studies. All the HoloLens 2 holograms presented in this paper have been tested with success by 30 in-service first responders in fire stations in Estonia by students (Guž, 2023; Liping, 2023; Pärnamets, 2023; Rasva, 2023). The aims of the bachelor theses, FightARs testing methodology as well as the results are published online in the Estonian language (Polikarpus et al., 2023). In our case study we first look at the dimensions of holograms, we then analyse whether the holograms are scalable or not and thirdly, we considered whether the trainee could interact with the hologram or not (see Figure 1).

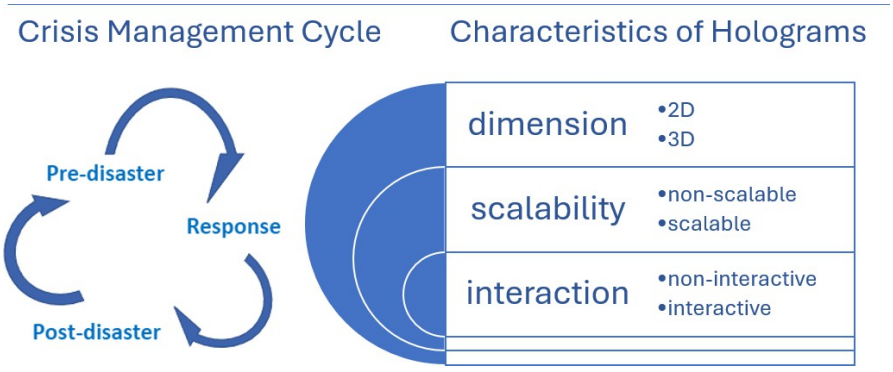


Figure 1. The characteristics of holograms and the crisis management cycle

AUGMENTED REALITY APPLICATION DESIGNED TO TRAIN FIREFIGHTERS: HOLOGRAM TYPES IN FIGHTARS

The FightARs application was developed within the FightARs project. HoloLens 2 augmented reality glasses

were used (Figure 2), where the holograms were tested and integrated in addition to which educational materials were also created. The topics of educational materials covered: first aid, situation awareness, UN marking of dangerous goods and deactivation of an electric car. The FightARs application and its code, QR-codes, voice commands' guide, pedagogical guidelines and FightARs best practice guide can be downloaded from the project homepage (http://fight-ar.com/) free of charge.



Figure 2. HoloLens 2 augmented reality classes and their use (Photo by Stella Polikarpus)

The application was connected to a learning management system (LMS) (Figure 3). Figure 3 shows the connection of 2D (the blue menu on the left) and 3D holograms (see in the centre of the figure). The project homepage and LMS are published on a WordPress web-based platform. 2D holograms were linked to the LMS specific pages in the FightARs course (an example can be seen in Figure 3 on the right). Connecting the AR application with a web-platform through 2D holograms meant that the learning materials could be automatically updated.

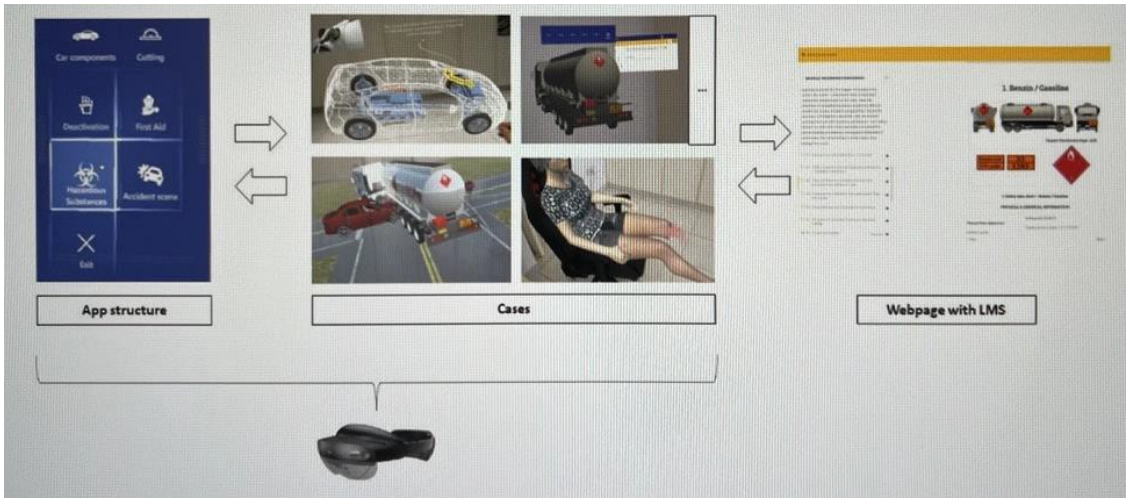


Figure 3. FightARs application connection to holograms and LMS (FightARs Partners 2023)

If the FightARs application is installed in HoloLens 2, the app icon could be found on the menu list (see Figure 4). In the first picture on the left we have illustrated the main menu of HoloLens 2. The right picture is the 3D hologram that pops up when FightARs application is “clicked” on in the main menu. The user needs to push the play button (Figure. 4 right) to open the 2D main menu of the FightARs application (Figure 5.).

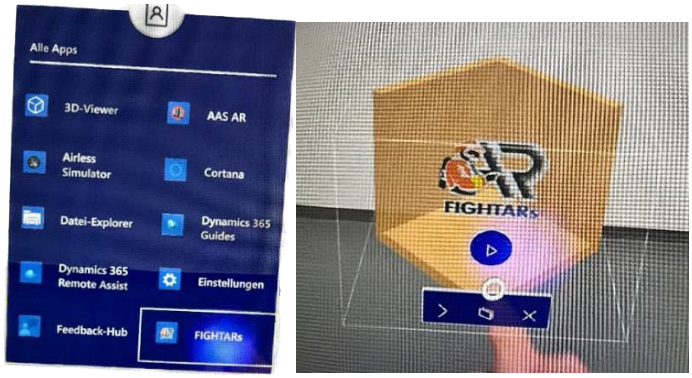


Figure 4. FightARs application icon in HoloLens 2 menu (FightARs Partners 2023)

The FightARs application menu gives access to all learning content and holograms created in the project. To access these designated sections, referred to as 'chapters,' users were required to select them from the main menu of the FightARs application (see Figure 5).

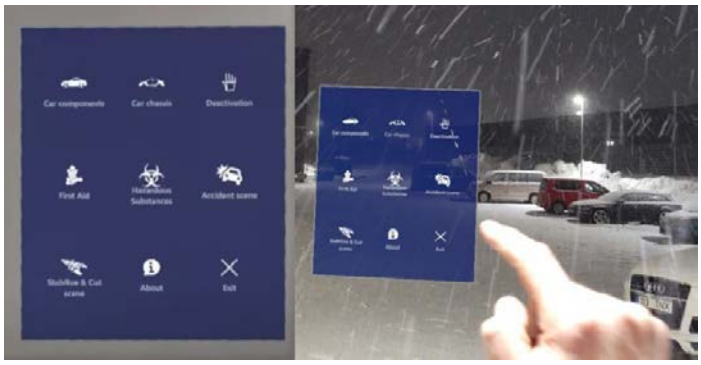


Figure 5. FightARs application menu (FightARs Partners 2023 and Liping, 2023)

First, we looked at the *dimensions of holograms*. In this paper, 2D holograms are holograms where the menu of the AR glasses is opened to move around in the application (e.g. Figure 5). Secondly, we identified if it is a *scalable hologram*, which implies that the user has the capability to independently modify the x, y, or z axis, or alternatively, adjust all axes simultaneously while maintaining consistent proportions (see the two upper photos from the case studies in Figure 3). If it is a non-scalable hologram, the user might be able to decide where in the physical room he/she wants the hologram to appear, but the user cannot reduce or increase its size (see the two lower photos indicating a woman and a collided truck in Figure 3).

The third important characteristic of a hologram is whether the user can interact with it or not. By *interaction*, we consider a hologram that the trainee can “click” on or in the future type something into the hologram and then the application checks the accuracy of the answer. Not every interaction with a hologram constitutes a digital twin, as the hologram may not simulate a real-life process. Instead, it could be designed to assess the correct sequence of the activities of the trainee (e.g. interaction in the form of a 2D hologram such as quizzes created using H5P). In FightARs two interactive holograms were used: car cutting and electrical car deactivation. These interactive 3D holograms were programmed inside the application to be responsive to the actions of the trainees and therefore can also be called digital twins. An overview of the dimensions, scalability, and interactivity of the FightAR's holograms and use cases are provided in Table 1.

Table 1. FightARs' hologram types, characteristics and use cases

Hologram types in FightARs	Dimen-sion	Scalable / non-scalable	Interactive / non-interactive	Description of use cases

2D hologram	2D	LMS window is scalable	Non-interactive or interactive e.g. H5P content	Used in the app to open LMS or application menus
3D non-interactive hologram	3D	Scalable	Non-interactive	Used to review firefighters' knowledge about UN hazard marking on a digital twin truck (Guž, 2023)
QR-code scan for 3D hologram	3D	Non-scalable	Non-interactive	Used to train triage of digital twin victims and first aid (Pärnamets, 2023)
Holographic scene	3D	Non-scalable	Non-interactive	Used to measure incident commanders' situation awareness with quantitative analyses of situation awareness on a digital twin accident site (Liping, 2023)
3D interactive hologram	3D	Scalable	Interactive	Used to train firefighters to respond to traffic accidents with digital twin electrical car (Rasva, 2023)

Non-Interactive Scalable Holograms

2D holograms linked to LMS in the FightARs app gave some freedom to trainers to tell a new story or give a new task with the same 3D hologram. This approach significantly enhances the efficiency of updating educational content, as it primarily involves modifications to the webpage, thereby reducing the necessity for extensive programming efforts. The only issue with this type of holograms is that HoloLens 2 needs to be connected to Wi-Fi, the connection must be stable, and the trainee must log onto the LMS. These 2D holograms could not be scaled, but they moved and some of them could be pinned in the air to fix their location in the physical room. As opposed to this, the windows of the LMS could be scaled. Rasva (2023, p. 33) notes that using the HoloLens 2 induced fatigue among users, suggesting that tasks typically performed on a LMS are more suitably executed on a flat screen device rather than reading lengthy text via 2D holograms associated with the LMS.

3D non-interactive scalable holograms were used to train firefighters' knowledge about the parts of an electrical car (see Figure 6) and for identifying the markings of trucks carrying hazardous material (see Figure 7).

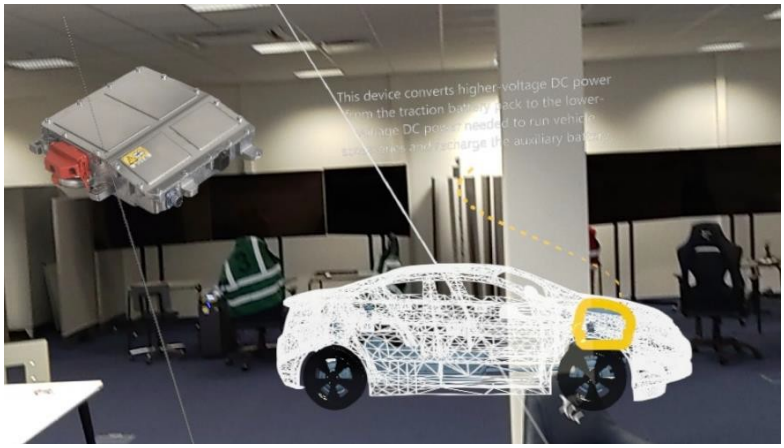


Figure 6. 3D digital twin scalable but non-interactive part of an electrical car

These holograms could be positioned, relocated, rotated anywhere within the room, and scaled to the authentic dimensions of a truck or passenger car. They were used to either gain new knowledge (Rasva, 2023) or review and assess existing knowledge and misconceptions trainees might have had (Guž, 2023).



Figure 7. 3D non-interactive hologram that can be scaled and moved (FightARs Partners 2023)

Non-Scalable Real-Size Holograms and QR-Code Scanning

QR-code scan to open 3D non-interactive hologram allowed to connect specific victim holograms to the physical training environment. In the study a variety of learning scenarios was created to train firefighters' and rescue incident commanders' first aid skills (Pärnamets, 2023). These holograms can be superimposed on dummies, allowing for the practical training of CPR skills, which is why they are non-scalable. These 5 holograms opened with 5 different QR-codes could be used simultaneously and create an accident with multiple casualties (see Figure 8). All these holograms were non-interactive holograms. However, web-based 2D LMS window was opened together with the hologram and their videos, H5P content or something else could be added for interactivity. It was interesting to open the hologram victim by scanning the QR-code in real life. After this, the scanning mode was kept active because sometimes the hologram could be moved in the physical room by moving the paper with the QR-code. If several holograms were opened at the same time, it was sometimes not possible for HoloLens 2 to process the situation well, after which the application crashed.



Figure 8. QR-code scan to open non-interactive 3D holograms. There are three victim holograms on the photo and above two victims head is LMS 2D hologram opened (FightARs Partners 2023)

Holographic scene (see Figure 9) is a term taken into use to describe non-interactive and non-scalable almost real-life sized 3D hologram that is a digital copy for the whole traffic accident scene (Liping, 2023). Navigation within a virtual reality environment is facilitated through the use of input devices such as joysticks, keyboards, gamepads, or other similar tools. Therefore, information collection from the VR scene for attaining situation awareness is different than in real life. To the contrary, a holographic scene allows to gather information similarly as in real life, through walking and getting new information. A holographic scene hologram can be overlayed with

real cars and then practical skills of cutting the car can be practiced. In Figure 9, (see the left picture) the holographic scene is opened in a mat room indoors and in the right picture (see Figure 9) it is opened outside, where the holographic red car with holographic victims is overlaying a real-life red car.



Figure 9. Holographic scene is 3-D is almost real-life size hologram (FightARs Partners 2023)

Interactive Scalable Holograms

A 3D interactive and scalable hologram is the most difficult digital twin type of hologram to program (see Figure 10). However, when this type of hologram is overlayed with real life objects, very interesting mixed-reality experiences can be created for users. In the application FightARs there are two types of this hologram, one is to deactivate an electrical car (upper picture in Figure 10) and the other is to select correct tools to rescue a victim from the passenger car (see the picture below in Figure 10). To use HoloLens 2 and interact with the holograms, trainees needed prior training on how to use HoloLens 2 (Rasva, 2023). Another issue emerging from that study is connected to language. Namely, when translating content in the LMS into all languages is rather easy, then creating the whole HoloLens 2 application in different languages is rather time consuming as it needs to be programmed in different languages and afterwards reinstalled.

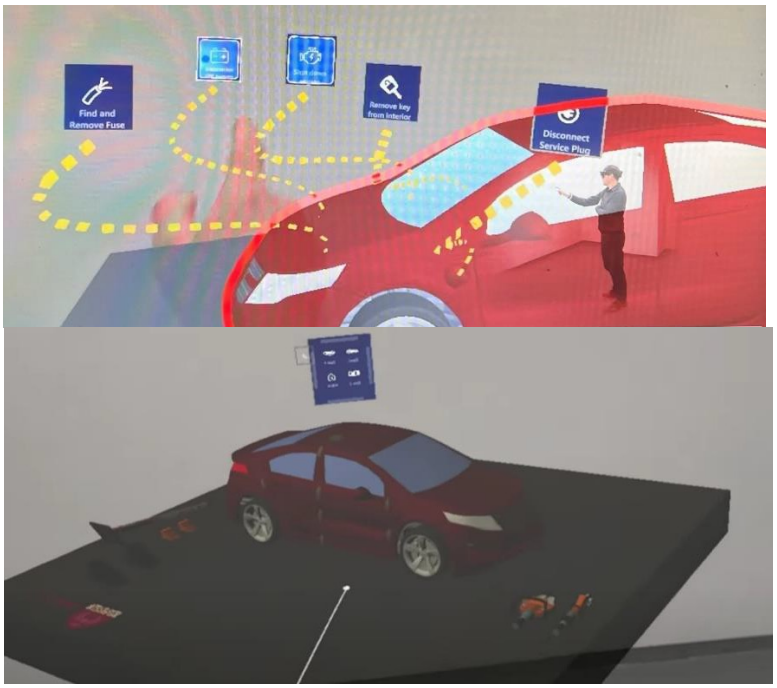


Figure 10. A 3D interactive and scalable hologram as a digital twin model created to interact with the actions of trainees (FightARs Partners 2023)

POTENTIAL USE OF AR IN CRISES MANAGEMENT TRAINING

We suggest that AR firstly needs to be taken into use during the pre-disaster phase. When firefighters and rescue incident commanders are accustomed to train with the AR application, they will have acquired practical skills to use HoloLens 2 like scaling and turning the holograms. If better internet connection were to be integrated into

incident response vehicles, we believe that AR devices could be used in the response phase in the future as well. For example, we believe that communication efficiency could be enhanced with HoloLens 2. A commercialised product Dynamics 365 Remote Assist was tested during the FightARs project and was found to be useful to show others on the screen what the person wearing HoloLens 2 is seeing (see Figure 11).

Needless to say, the utilization of this tool necessitates the acquisition of a licensed program. Nevertheless, we see a big potential for AR and HTC in the response phase of crises, because the "caller" can have their hands free for emergency actions and the "adviser" can see the same visual as the caller. Simultaneously, the advisor is visible for the caller as a 2D hologram and clear audio is heard from the AR-glasses. In addition to this, calls from AR-glasses to AR-glasses are an option.



Figure 11. A trainee using HoloLens 2 vision is seen on screen in class through the Dynamics 365 Remote Assist software. HoloLens 2 could also be connected to screens using mirroring functions.

Post-disaster phase is about learning from real incidents. In the post-disaster phase, feedback from experiences is analyzed to improve the anticipation and forecasting of potential issues, identify risks, and revise established procedures (Aligne & Mattioli, 2011). Naturally, if AR is employed during the response phase, its usage should extend to the post-disaster phase, or at a minimum, experiences from its application should be gathered and the employed procedures, ethics, and equipment critically analyzed. In Table 2 we have brought together the crises management phases and studied hologram types.

Table 2. The connection of FightARs hologram types to crises management phases

Hologram types in FightARs	Phase	Examples	Knowledge and Skills
2D hologram	All 3 phases	Video calls, representation of sensor data, resource allocation, AI decision suggestions, recordings, photos, LMS and H5P in it, guides in pdf files etc.	Standard operational procedures or details of an equipment, or any other text or 2D visual information.
3D non-interactive hologram	In pre-disaster and response phases	Examples Figure 6 and 7. In response phase points of interests (POI) could be marked (Christaki et al., 2022).	To enhance knowledge of firefighters with 3D objects when composition of the object needs to be well understood or POIs marked in learning environment.
QR-code scan for 3D hologram	In pre-disaster and response phases	Example Figure 8. In response phase POIs could be marked (Christaki et al., 2022) using QR-codes and then information could be updated and read from the web.	To create any practical skills training scenario in training ground or to guide self-directed learning
Holographic scene	In pre-disaster phase	Example Figure 9.	Same as above, plus dynamic decision-making training
3D interactive hologram	In pre-disaster phase	Example Figure 10. digital twin.	Better to use VR or LMS, because skill to use HoloLens is need.

DISCUSSION AND FUTURE RESEARCH

AR has the potential to improve situation awareness and through that dynamic the decision-making of first

responders (Christaki et al., 2022) and firefighters while smoke-diving (Gkika et al., 2023). However, to train staff to execute safe actions during incidents, the implementation of new technologies should start from the pre-disaster phase. We analysed AR digital holograms in the training of first responders. We wanted to find out what type of holograms support firefighters' knowledge and skills training in which crisis management phase. For this reason, we analysed three hologram characteristics: dimension, scalability and interactivity in the application FightARs.

In Figure 12 we present the hierarchy of these digital holograms. 3D non-interactive scalable holograms or 3D non-scalable real-size holograms like a holographic scene seem to be the best option to train practical skills, because they can be overlayed with physical objects, or the physical environment can be enriched with holograms using QR-codes. Besides training practical skills like cutting a car or sealing a leak, a holographic scene is a good way to train situation awareness and dynamic decision-making. However, programming these 3D models for AR is rather time consuming compared to VR. 3D interactive scalable holograms or digital twins are the hardest to program and to use. To use these holograms, trainees and trainers first need to learn the skill to use HoloLens 2, they need to get used to the specific movements required for each corresponding action and related processes. These practical skills should be learned by first responders during the pre-disaster phase only if AR technologies will be used in the response phase in the future. Otherwise, it merely consumes valuable training time.

We know from the FightARs project that HoloLens 2 offers a possibility to make video calls (Figure 11), record videos, take photos, scan or place holograms etc. All these new functionalities should be further tested in order to find out how they can support situation awareness and dynamic decision-making and other practical skills training for firefighters. If HoloLens 2 were to be combined with a powerful laptop and great Wi-Fi connection, we could even use AI to help make predictions from sensor data collected by robots or firefighters about gas-cloud movements or fire spread. AR devices in combination with other technologies have great potential in all crises management phases.

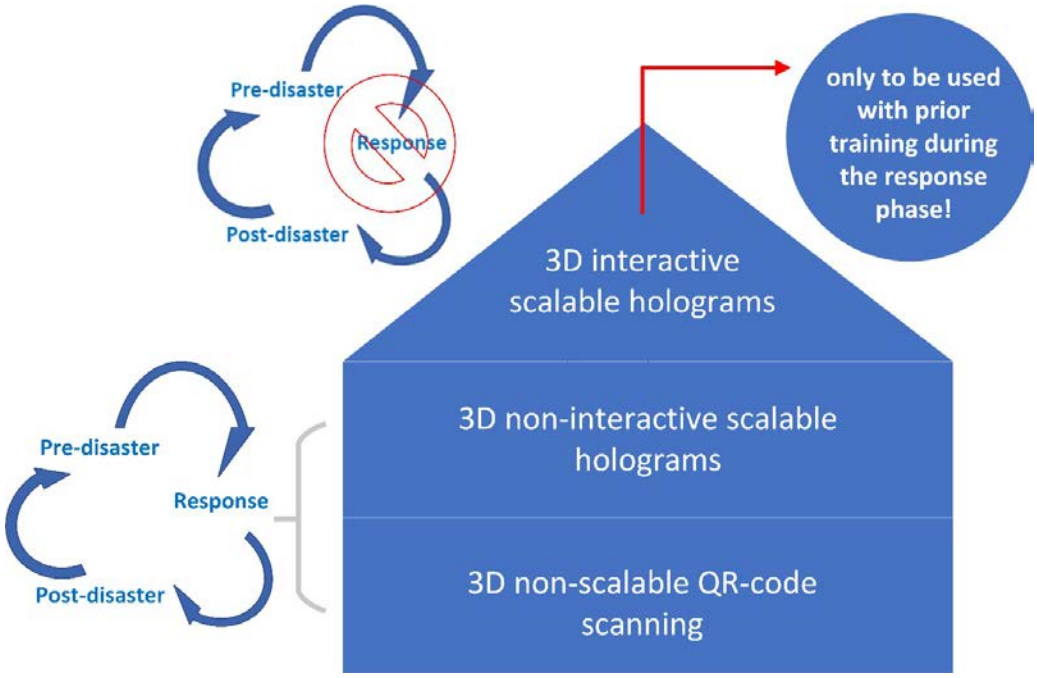


Figure 12. The hierarchy of holograms and crises management cycles

In this article we focused on the ability of HoloLens 2 to produce different holograms. From the experience of implementing VR in first responders' training, we claim that AR will be quickly implemented into firefighters' training if trainers can create 3D non-interactive scalable or non-scalable holograms themselves with the help of VR. VR has been used to train firefighters for a long time, therefore these 3D models already exist in VR. In VR, these models can be combined into holographic scenes and training scenarios (Polikarpus et al., 2021). If the scene or even scenario could be imported into HoloLens 2, it could enhance the research and training of situation

awareness to be more authentic and accessible. We could also train firefighters' practical skills when holograms are overlayed with reusable physical simulators that allow for example the use of hydraulic tools. Simultaneously, a virtual layer might allow interactions and information presentation that is not possible in the physical world.

We are a bit more hesitant about the need to develop interactive holograms by trainers themselves for AR as movements to interact with these holograms require more practical training on how to use HoloLens 2. Additionally, we are not sure if this practical skill of using AR is later needed in the response phase. Earlier research has stated that HoloLens 2 applications required some practice, especially by the majority of participants who had no prior experience with holographic displays (Christaki et al., 2022). We found out that problems with using AR solely appeared when interactive holograms were used, or if the trainee had to scale the hologram him/herself. A holographic scene and QR-code scan to open non-interactive holograms did not require any special skills from the trainee and therefore could these hologram types be more easily integrated into training and the response phase of crises. During the response phase, everything is time critical and has to work flawlessly. Therefore, if AR were to be used in the response phase, then all these types of holograms should be used in the pre-disaster phase in order to train the needed AR practical skills. In addition to using AR in training, the technology needs to be developed to be weatherproof, and not cause explosions or any other hazard for anyone.

Limitations

There are some limitations in our typology of holograms. The term interactivity can be misleading, as clicking on the menu does instigate a new hologram or menu, but we named these 2D holograms as non-interactive. Throughout the testing phase conducted in fire stations it was observed that the holographic scene, which is designed to be non-scalable and of authentic size, encountered spatial constraints within indoor environments. Consequently, its application was occasionally extended to outdoor settings. However, it is pertinent to note that the HoloLens technology is not inherently designed for outdoor operation. Also, the holographic scene had 15 possible holograms and in one case, there was a leak under the truck and in another case the smoke to make the scene more dynamic. Our typology did not discuss whether holograms had moving parts in 3D models. In other studies, they have been differentiated as dynamic and static holograms (Polikarpus, Terep, et al., 2023).

CONCLUSION

The study investigated what type of holograms support firefighters' knowledge and skills training in which crisis management phase. We propose that 3D non-interactive scalable or non-scalable holograms should be used to exercise practical firefighting skills. Mixed-reality experiences and holographic scenes should be created in VR by trainers themselves in future AR applications. The technology should enable being exported from VR and imported to AR to give firefighters and incident commanders a rather cheap, repeatable, and comparable way to train their situation awareness, dynamic decision-making, and other practical skills.

It is important to remember that any new technology used by first responders should improve their situational awareness and dynamic decision-making and end-users should have training and vocabulary to use and develop it. The technology ought to be flawless, easy to use and before being integrated into the response phase of crises, it first needs to be tested and used extensively in the pre-disaster phase during the training of first responders. This foundational integration in training contexts is essential for ensuring an effective utilization in future emergency situations and post-disaster analyses.

ACKNOWLEDGMENTS

We thank all members of the FightARs project and the students who conducted their final papers to test the application in firefighters' authentic training environments. Your contributions were essential for the completion of the project.

REFERENCES

Ahmad, A. S., Alomaier, A. T., Elmahal, D. M., Abdlfatah, R. F., & Ibrahim, D. M. (2021). EduGram: Education Development Based on Hologram Technology. *International Journal of Online and Biomedical Engineering (IJOE)*, *17*(14), 32–49. https://doi.org/10.3991/ijoe.v17i14.27371

Aligne, F., & Mattioli, J. (2011). *The Role of Context for Crisis Management Cycle* (pp. 113–132). https://doi.org/10.1007/978-1-4419-7406-8_6

Botín-Sanabria, D. M., Mihaita, A.-S., Peimbert-García, R. E., Ramírez-Moreno, M. A., Ramírez-Mendoza, R. A., & Lozoya-Santos, J. de J. (2022). Digital Twin Technology Challenges and Applications: A Comprehensive Review. *Remote Sensing*, *14*(6), 1335. https://doi.org/10.3390/rs14061335

Bullock, J., Haddow, George, & Coppola, Damon P. (2013). Chapter 10 Mitigation, Prevention, and Preparedness. In *Introduction to Homeland Security* (pp. 435–494). Elsevier. https://www.google.ee/books/edition/Introduction_to_Homeland_Security/s7COURn3XhMC?hl=et&gbpv=1&dq=Chapter+10+•+Mitigation,+Prevention,+and+Preparedness&pg=PA453&printsec=frontcover

Christaki, K., Tsiakmakis, D., Babic, I., Inglese, G., Konstantoudakis, K., Giunta, G., Dimou, A., Balet, O., & Daras, P. (2022). Augmented reality points of interest for improved first responder situational awareness. *Proceedings of the International ISCRAM Conference*, 755–770.

Elmahal, D. M., Ahmad, A. S., Alomaier, A. T., Abdlfatah, R. F., & Hussein, D. M. (2020). Comparative study between hologram technology and augmented reality. *Journal of Information Technology Management*, *12*(2), 90–106. https://doi.org/10.22059/JITM.2020.75794

FightARs. (2022). *FightARs – Augmented Reality for Firefighters*. http://fight-ar.com/

Gkika, I., Pattas, D., Konstantoudakis, K., & Zarpalas, D. (2023). Object Detection and Augmented Reality Annotations for Increased Situational Awareness in Light Smoke Conditions. *Proceedings of the International ISCRAM Conference*, 231–241. https://doi.org/10.59297/YOMA9043

Guž, V. (2023). *Ohtlike veoste tähistuste tuvastamine liitreaalsuses meeskonnavanemate näitel* [Sisekaitseakadeemia]. https://digiriiul.sisekaitse.ee/handle/123456789/3059

He, L., & Zhu, S. (2022). Virtual Reality Technology in Visual Design of Artistic Images: Analysis and Applications. *Scientific Programming*, *2022*, 1–6. https://doi.org/10.1155/2022/2527623

Heldal, I., & Wijkmark, C. H. (2017). Simulations and Serious Games for Firefighter Training: Users' Perspective. *14th ISCRAM Conference - Albi, France, May 2017*, *May*, 868–878. http://idl.iscram.org/files/ilonaheldal/2017/1518_IlonaHeldal+CeciliaHammarWijkmark2017.pdf

Herrington, J., & Oliver, R. (1995). Critical characteristics of situated learning: Implications for the instructional design of multimedia. In A. S. for C. in L. in T. E. A. 95 (Ed.), *ASCILITE -CONFERENCE PROCEEDINGS* (pp. 253–262). The university of Melbourne, Faculty of Science. http://www.konstruktivismus.uni-koeln.de/didaktik/situierteslernen/herrington.pdf%0Apapers3://publication/uuid/CE9062FA-63D0-44BC-BF6E-01313702ADC1

Hofmann, J., & Polikarpus, S. (2022). Application of Augmented Reality in firefighters training: From Safe to SafAR. In R. Wechner, M. Bues, & U. Kloos (Eds.), *GI VR / AR Workshop 2022* (pp. 1–9). Gesellschaft für Informatik. https://doi.org/10.18420/vrar2022_184

Işik, V. (2014). Classification of Holograms and Types of Hologram Used in Holographic Art. *Online Journal of Art and Design*, *2*(3), 15–26. http://www.adjournal.net/articles/23/232.pdf

Keskitalo, T. (2022). Pedagogical Practices for Organizing Simulation-Based Healthcare Education. *International Journal of Learning, Teaching and Educational Research*, *21*(4), 80–96. https://doi.org/10.26803/ijlter.21.4.1

Kurt, S. (2019). TPACK: Technological Pedagogical Content Knowledge Framework - Educational Technology. *Educational Technology*, 2023. https://educationaltechnology.net/technological-pedagogical-content-knowledge-tpack-framework/

Lamb, K., Farrow, M., Olymbios, C., Launder, D., & Greatbatch, I. (2021). Systematic incident command training and organisational competence. *International Journal of Emergency Services*, *10*(2), 222–234. https://doi.org/10.1108/IJES-05-2020-0029

Liping, L. (2023). *Holograafilise sündmuskoha rakendamise võimalused päästemeeskonna juhtide olukorrateadlikkuse arendamisel* [Sisekaitaseakadeemia]. https://digiriiul.sisekaitse.ee/handle/123456789/3064

Ma, J. Y., & Choi, J. S. (2007). The Virtuality and Reality of Augmented Reality. *Journal of Multimedia*, *2*(1),

32–37. https://doi.org/10.4304/jmm.2.1.32-37

Pärnamets, J. M. (2023). *Liitreaalsuse kasutamine esmaabi õppes päästekomandodes* [Sisekaitseakadeemia]. https://digiriiul.sisekaitse.ee/handle/123456789/3057

Polikarpus, S., Bøhm, M., & Ley, T. (2019). Training Incident Commander's Situational Awareness—A Discussion of How Simulation Software Facilitate Learning. In T. Väljataga & M. Laanpere (Eds.), *Digital Turn in Schools—Research, Policy, Practice* (pp. 219–234). Springer, Singapore. https://doi.org/10.1007/978-981-13-7361-9_15

Polikarpus, S., Ley, T., & Poom-Valickis, K. (2021). Collaborative Authoring of Virtual Simulation Scenarios for Assessing Situational Awareness. In A. Adrot, R. Grace, K. Moore, & C. Zobel (Eds.), *Proceedings of the 18th ISCRAM Conference* (pp. 229–237). Blacksburg, VA, USA. https://www.drrm.fralinlifesci.vt.edu/iscram2021/files/ISCRAM2021Proceedings.pdf

Polikarpus, S., Luik, P., Poom-Valickis, K., & Ley, T. (2023). The Role of Trainers in Implementing Virtual Simulation-based Training: Effects on Attitude and TPACK Knowledge. *Vocations and Learning*, *16*(3), 459–486. https://doi.org/10.1007/s12186-023-09322-1

Polikarpus, S., Sarmiento-Márquez, E. M., & Ley, T. (2023). Creation and Use of Virtual Simulations for Measuring Situation Awareness of Incident Commanders. In T. Gjøsæter, J. Radianti, & Y. Murayama (Eds.), *ITDRR2022* (Informatio, pp. 23–38). Springer, Cham. https://doi.org/10.1007/978-3-031-34207-3_2

Polikarpus, S., Terep, T., & Reitel, K. (2023). Liitreaalsuse kasutamine väljaõppes: pääste tulevik? *Turvalisuskompass*, *5*(2), 235–267. https://doi.org/10.15158/W11G-MP89

Rasva, A. (2023). *Liitreaalsuse kasutamine elektriautode avariide lahendamise koolitamisel päästeteenistujate väljaõppes* [Sisekaitseakadeemia]. https://digiriiul.sisekaitse.ee/handle/123456789/3056

Rauschnabel, P. A., Brem, A., & Ro, Y. K. (2015). Augmented Reality Smart Glasses : Definition, Conceptual Insights, and Managerial Importance. In -. https://www.researchgate.net/profile/Alexander-Brem-2/publication/279942768_Augmented_Reality_Smart_Glasses_Definition_Conceptual_Insights_and_Managerial_Importance/links/5721ec2e08aee857c3b5dd6c/Augmented-Reality-Smart-Glasses-Definition-Conceptual-Insig

Reisoğlu, I., Topu, B., Yılmaz, R., Karakuş Yılmaz, T., & Göktaş, Y. (2017). 3D virtual learning environments in education: a meta-review. *Asia Pacific Education Review*, *18*(1), 81–100. https://doi.org/10.1007/s12564-016-9467-0

The Fightars Partnership. (2021). *FihtARs Pedagogical Guidelines* (p. 46). http://fight-ar.com/

Wijkmark, C. H., Bail, R. D. F., & Metallinou, M. (2022). Introducing Virtual Reality for Firefighter Skills Training Opinions from Sweden and Brazil. *Journal of Applied Technical and Educational Sciences JATES*, *12*(3), 1–24. https://doi.org/10.24368/jates.v12iY.Z

Wijkmark, C. H., Heldal, I., & Metallinou, M.-M. (2021). Experiencing Immersive VR Simulation for Firefighter Skills Training. In A. Adrot, R. Grace, M. Kathleen, & C. Zobel (Eds.), *Proceedings of the 18th ISCRAM Conference* (Issue May, pp. 913–921). http://idl.iscram.org/files/ceciliahammarwijkmark/2021/2383_CeciliaHammarWijkmark_etal2021.pdf

Yoo, H., Jang, J., Oh, H., & Park, I. (2022). The potentials and trends of holography in education: A scoping review. *Computers & Education*, *186*(January), 104533. https://doi.org/10.1016/j.compedu.2022.104533



Information Systems in Humanitarian Operations

Track Chairs

Amin Maghsoudi, Hanken School of Economics

Christian Kalla, University of Münster

Hossein Baharmand, University of Agder

Navigating the Digital Technology Behavior for the Distribution of Humanitarian Cash-Based Assistance: Aid Recipients Experience

Amin Maghsoudi
HUMLOG Institute, Hanken School of Economics, Helsinki, Finland
Amin.maghsoudi@hanken.fi

Abdelsalam Hamid Abakar
Department of Management, College of Business Administration, A'Sharqiyah University, Ibra, Oman;
HUMLOG Institute, Hanken School of Economics, Helsinki, Finland
abdelsalam.hamidabakar@hanken.fi

ABSTRACT

Cash-based assistance (CBA) programs have become an increasingly common approach as one modality for the provision of humanitarian aid. The use of disruptive and digital technologies (such as mobile money, electronic vouchers) has provided benefits to implementation of such programs in terms of for instance, cost effectiveness, timeliness, coverage, safety, and financial inclusion. Despite the benefits, the digitalization of such programs is very context-dependent, including the needs and characteristics of affected populations as end recipients of the assistance. This research is aimed to investigate the contextual characteristics toward digital transformation addressing the end user's experience. The series of interviews is conducted from humanitarian practitioners based in Uganda. The findings are yet to be analyzed, but some initial findings are discussed, and future research is addressed in the conclusion.

Keywords

Cash-based assistance, digital transition, adoption behavior, refugee camps, aid recipients.

INTRODUCTION

The global crisis is rapidly changing and growing with the lethal mix of conflict (e.g., Sudan, Gaza, Ukraine, Syria), COVID-19 pandemic, food insecurity, and climate emergency. The number of forcibly displaced people due to such crises has continued to increase from 82.4 million in 2020 to more than 103 million, of which 32,5 million refugees, by the end of 2022 (UNHCR, 2023). According to the EU, more than 35% of the humanitarian assistance reaches refugees in the form of cash-based assistance (CBA) (e.g., via different delivery mechanism such as physical cash, mobile money, debit card) (ECHO, 2021) and to a larger scale of cash programs above 10 million euro (DG-ECHO, 2024) . DG-ECHO emphasizes the links between cash and the secure and efficient use of digital tools. The EU promotes cash operations, which build on technological advances and innovations and use technology to strengthen interoperability and link to longer-term solutions. EU humanitarian aid prefers digital solutions such as electronic cash transfers and digital identities. Despite tremendous upward trend toward delivery of CBA programs, the effectiveness and cost-efficiency of CBA implementation is highly dependent on the contextual factors such as the financial and technological infrastructure, regulatory framework, together with market, economic conditions of the host countries, and more importantly the contextual characteristics of a id recipients (Heaslip et al., 2018b; Kian et al., 2022; Piotrowicz, 2018; Wilson et al., 2018).

WiP Paper – Information Systems in Humanitarian Operations
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

Humanitarian organizations (HOs) have been increasingly using digital technologies along the cash operations cycle from needs assessment to the cash transfers via different delivery mechanism such as mobile money, bank accounts or smart card (Jodar, 2021). Digitalization has been implemented and adapted in different ways and contexts with a typical aim of improving access to humanitarian a id. For instance, in Sudan, there has been a shift from emergency food assistance to digital cash transfers. In India, the public distribution system is intervened through digital transformation since 2010, and while in the UK, welfare has been digital since 2012 and from 2016, the humanitarian assistance for asylum seekers is provided through biometrics and debit cards (Jaspars & Sathyamala, 2021). Although using technology for humanitarian CBA is nothing new, the phenomena are relatively emerging, at least in academic literature. In this context, "digital CBA" refers to the "transfer of money from the implementing organization or its delegates to program beneficiaries via electronic means (e.g., mobile, card, or account). That indicates both the financial service provider (FSP) and beneficiary use an electronic medium of transfers (CaLP, 2013).

While aid recipients at the downstream supply chain are at the center of decision -making process, their characteristics must be properly determined to ensure the principle of "do no digital harm" is upheld (Burton, 2020). Digitalized CBA programs are receiving significant attention from scholars, while the need and role of a id recipients as a critical stakeholder in the humanitarian aid setting is understudied. Unlike CBA program, the success of digitalized CBA depends highly on the aid recipient's choice, inclusion, divide and so many others (IOM, 2020; Patil et al., 2020). For instance, factors like digital literacy of the beneficiary, phone ownership, access to internet, ability to use internet, having digital ID, can be detrimental for the success of digitalization of CBA in the humanitarian contexts (Mitchell et al., 2019). One of the main criticisms against such an approach is that it is often exclusionary (includes only those who can use digital technology) and goes against the ethos of humanitarian activities (Burton, 2020) . Specifically, such digital transition is extremely challenging in the response to refugees in disaster-prone areas such as Uganda, Congo, or Sudan. Among those, some studies have found non-significant relationship however, as consumers consider ease of using mobile as unimportant (Aslam et al., 2017). In other cases, aid recipients with long queuing times with mobile service providers have made their complaints over the cash receipt services, and thus perceived the complexity and extra efforts needed for such digital services in Ethiopia (Murray, 2016). Despite, large number of populations receiving CBA through all means of digital, the digital transition varies from country to country particularly in the context of Africa (Burton, 2020; Maghsoudi, Harpring, Piotrowicz, & Kedziora, 2023).

In this study, we aim to investigate and assess the contextual factors that are significant for the success of digitalization of CBA from a beneficiary's or aid recipients' perspective. We rely in interviews conducted in refugee camps in Uganda, minutes of meeting, and several published literatures and reports available in the public domain to build the framework which places beneficiaries at the core of digitalized CBA programs. The results are under work-in-progress.

BACKGROUND

This section includes the summary of findings obtained based on the literature review. Overall, implementing the digital cash-based assistance program in the humanitarian aid sector is complicated and challenged by many factors and surrounding circumstances, including beneficiaries, the country hosting the refugees, the local community, the type and objectives of the assistance, and the partners involved directly or indirectly. The study and the findings are adapted the TELOS framework for assessing the fea sibility of digital CBA programs by focusing on the societal aspects. In this section, we found that the contextual factors of digital CBA are in line with the TELOS framework from a broader perspective. TELOS framework refers to five areas of feasibility criteria (i.e., Technical, Economic, Legal, Operational, Societal). TELOS framework helps researchers to determine an optimal solution to evaluating the level of risks, costs, and benefits of alternatives/options based on several feasibility areas (Drljaca & Latinovic, 2018). The TELOS feasibility model assesses the potential for implementation of a new technology. This framework is used in many fields, e.g., engineering and technology, and has also been applied to medicine (Segawa & Muzinda, 2021). After reviewing HO re ports, and recommendations from digital CBA programs, and aligning the reports' output with TELOS framework, we provide a conceptual evaluation framework including the societal factors obtained from the findings. The finding is synthesized in this section to provide and grasp an overview of factors to consider for the design and implementation of digital CBA in the refugee context.

Numerous factors associated with digital CBA have been reported. However, the context-related factors cannot be comprehended at one point in time since these factors vary based on each case. Such contextual factors are associated with the implementation context's general conditions, such as infrastructures, host country's regulations, technology level, and availability at the program place (Abdelgawad et al., 2023; Juntunen et al., 2023; Maghsoudi, Harpring, Piotrowicz, & Heaslip, 2023; Maghsoudi, Harpring, Piotrowicz, & Kedziora, 2023).

The other association is with program-specific related conditions such, as the specific operational requirements (banking or financial systems and regulations, specific vendors, digital financial service providers, agents networks), besides related societa l factors such as financial literacy, digital literacy, and gender inclusion (Burton, 2020; Heaslip et al., 2018; Tappis & Doocy, 2018).

There are many societal and behavioral issues associated with digital cash assistance, based on what humanitarian organizations have reported. Societal factors can be described as the society and beneficiaries' willingness and acceptance of technology (UNHCR, 2020). These experiences revealed topics such as the long-term socio-economic impact (e.g., beneficiaries' preference, technology acceptance, financial inclusion of digital cash assistance, mobile phone ownership, discrimination in delivering digital cash-based assistance, local community language issues) as well as the responsiveness of digital cash mechanism to the beneficiaries' complaints (CWG, 2020). In addition to the target beneficiaries, some issues are closely associated with the program target beneficiaries (e.g., mobile ownership, SIM card, IDs, and costs of mobile and internet service) (Abdelgawad et al., 2023; Maciej Serda et al., 2020). Furthermore, there are some emerging issues (e.g., beneficiaries' preferences, ease of use, complaint channels, and local language barriers) (Maghsoudi et al., 2023; Jontunen et al., 2023). Since the last decade, the digitalization has been implemented and adapted in different ways and contexts with a typical aim of improving access to humanitarian assistance. For instance, in Sudan, there has been a shift from emergency food assistance to digital CBA, through the establishment of a new national cash-based family support program. The use of digital CBA is even accelerated during the pandemic due to the social-distancing measures and its called for the increased digitalization of humanitarian CBA pushing toward more remote programming and digital delivery mechanisms (Lawson-McDowall et al., 2021).

In the same vein, societal factors are crucial in digital cash-based assistance. From a societal perspective, several issues were identified, such as literacy both technological and financial. In refugees setting the illiterate rate, lack of technological skills, besides lack of recognized ID or documents, all are factors that characterize a refugees community or beneficiaries community in emerging economies (Doocy et al., 2011). The significant factors also beneficiaries related such as the cost of having a mobile phone, smart or appropriate phone due to income issues and the small amount of money that refugees received (CWG, 2020). Owning a smartphone is a challenge for refugees in Uganda and other contexts. Other factors that challenge refugees are acceptance and preference, which represents how the program is perceived and accepted by refugees and the ease of use of the program, particularly for illiterate people and people without technology experience, gender inequality is a big issue, especially for women. Women are more vulnerable because a considerable percentage, around 45 percent of women in low-income countries, do not have IDs. Therefore, they cannot apply for any official transaction, open a bank account, or own a SIM Card since they need IDs and proper documents. All the above factors and other sub-factors characterize the implementation of digital CBA.

RESEARCH METHOD

The research method is qualitative in nature using interviews conducted with humanitarian practitioners based in Uganda. This first round of analysis focuses on interviews conducted with humanitarian organizations (INGOs, UN agencies) responding to the refugee crisis in Uganda. The responded organizations were part of local cash working groups in Uganda covering 14 refugee camps in different locations across the country, and the data collection and interview process was led by one of the project partners loca lly based in Uganda in 2022. Reference group members also assisted the project partners for the data collection process. The interviewees are selected with purposeful sample selection based on their experience in CBA program, availability and accessibility given their contribution in terms of CBA in response to the refugee crisis in Uganda providing different assistance from food, to health, protection and sanitation support. The organizations are included international NGOs, UN agencies, and those with representatives at the local field level with the direct link serving beneficiaries or through their local implementing partners. The interviewee positions are varied from head of CBA program, to community service officer, team leaders, project coordinator and market officer given the extensive knowledge of CBA programs implementation. Most of the interviewees are from top to mid-level manger.

Uganda Case

Uganda was selected for the purpose of study as one-fifth of its population are is still estimated to live under the national poverty line. The backbone of the economy is the agriculture sector with primarily smallholder farmers. Access to productive land varies significantly on the prevalent land tenure system, the country's rapidly growing population is putting pressure on land resources, and the allocation of land to refugees becoming m ore difficult. With over 1.5 million registered refugees, Uganda is the third-largest hosting country in the world (and the largest in Africa)(UNHCR, 2021).

Cash-based assistance in Uganda has a long history. Since 2019, the number of refugees has stayed around 3

million, with occasional increases. Since 2019, nearly 40% of cash transfers have taken place through either mobile money (31.1%) or through bank accounts (9.8%). Other modalities, i.e., pre-paid cards, e-vouchers, and paper cards make just about 1% of the total cash transfer value. Another major observation from Uganda is essentially the nature of cash-based interventions and their modality. About 80% and 61% of cash -based interventions were used for food security in the years 2021 and 2022 (till June) respectively. Almost the entire cash intervention for food security is being handled through the World Food Programme (WFP).

This has been observed by the Uganda Cash Consortium (UCC), a consortium of international NGOs operating out of Uganda, and have made a three-year plan for developing a Multipurpose cash assistance (MPCT) program to make CBA a more inclusive mechanism for a wide variety of assistance programs. UCC identifies that for the successful working of MPCT, they need to focus on a program agenda that includes financial literacy training for beneficiaries, linking long-term livelihood assistance programs with CBA and promoting digital payments, and increasing digital cash transfers. The first phase of MPCT concluded in July 2022 and the second phase is currently underway. Phase 2 focuses on targeting the beneficiaries and linking their needs with the CBA. To this end, UCC focuses on three vulnerability categories, i.e., protection-specific vulnerability, economic-specific vulnerability, and sector-specific vulnerability.

However, it is important to note that, there are certain structural challenges to operationalizing CBA in Uganda in the current form. One most critical challenge are the stark difference in the choice of modality between the financial service providers (FSP) and the beneficiaries. While FSPs overwhelmingly prefer digital transactions, beneficiaries prefer hard cash. This is evidenced by the fact that more than six billion UGX has been transferred through banking institutions in 2019-2020 while less than one billion UGX has been transferred through mobile network operators and microfinance institutions during this period. On the other hand, only 4% of the beneficiaries prefer transfers through banks and the majority (70%) prefer it through mobile money. Some of the arguments form beneficiaries for this preference include ease of access and lack of transport required for beneficiaries, the presence of agents inside the settlement means beneficiaries do not have to travel far from their homes to open accounts and make withdrawals. Hence, beneficiaries see mobile money as more inclusive for people who have trouble traveling long distances, such as older persons, pregnant women, or persons with disabilities. Opening a mobile money account was a relatively simple process for beneficiaries to understand the costs associated with opening and maintaining the account were limited and mobile money is confidential and discreet, especially compared to the highly visible distributions of direct or OTC cash (REACH, 2021).

PRELIMINARY FINDINGS FROM INTERVIEWS

A total of 8 interviews were conducted with humanitarian organizations. The interviews were conducted and transcribed and analyzed in NVivo software. The qualitative framework that was presented divided the interview data into four distinct categories of actors, infrastructure, beneficiaries and financial assistance following the TELOS framework features for the societal aspects. The interviews were transcribed and coded. The coding process was started through initial coding (Saldaña, 2013), where the resea rcher familiarises themselves with the data, breaks it down into parts, examines differences and similarities (e.g., between different stakeholders), and makes notes on the data. Initial coding generally produces a set of topics, under which more specific codes can be placed. The 1st order concepts are themes that emerge from the text. In the data structure, they are summarized significantly, as the number of these themes in the original coding is large. They are then organized into 2 nd order concepts according to common denominators. The aggregate dimensions are determined through the research objectives as well as the first two orders in the data structure. The findings from the interviews are organized according to this structure. The analysis section of the findings is work-in-progress.

The roles of actors, infrastructure and environment, and elements of financial assistance are factors that can be analyzed quite effectively through scientific literature, secondary data, and the interviews already conducted. However, arguably the most important stakeholders', the beneficiaries' perspective remains difficult to ascertain. Through the interviews with stakeholders who regularly work with beneficiaries, we were able to establish the most relevant underlying factors that contribute to successful CBA programs by means of digital technologies.

Actors within these programs and the partnerships that they form in order to design, implement, and manage them are central to the success. This was apparent in the workshop, where participants represented HOs and financial institutions. The workshop was conducted to validate the results. The workshop was organized for the stakeholders who participated in the stakeholder interviews conducted in 2022. The workshop was a success as a total of 9 partners from 8 organizations were present. The workshop took place in Uganda and was also streamed online to give, for example, partners in Europe a chance to participate. It was structured so that after a round of introductions. The workshop then proceeded to the discussion, which was divided into four themes: startin g point

for DigCBA, gender disparities, access to services, and cash-based assistance. These themes were derived from the literature as well as the interviews findings. Before giving the floor to the workshop participants, an audience engagement tool Wooclap was used to gather quantitative data and give some background to the discussion. The workshop with the presentations and discussion lasted two and a half hours and was considered useful by both the participants and the organizers Section 2.2 presents the quantitative results of the Wooclap, along with some additional qualitative analysis. The workshop were held about three hours and with a break in between.

The infrastructure and environment ties in closely with the actors and partnerships, as for example the government is responsible for the regulations regarding refugees. This has significant relevance to the integration of refugees into societies, and whether they reside in camps or urban areas. CBA requires specific contextual elements, such as a functioning market where the cash can be used. Therefore, the type of assistance administered needs to be in line with the infrastructure. The most important factors, however remain the preferences and capabilities of the beneficiaries. The issue of financial literacy was raised, not just in regard to the beneficiaries, but aid workers as well working within the CBA field. Financial literacy is a challenging aspect, as refugees often need to deal with multiple currencies and modes of payment. CBA is also done in collaboration with financial institutions, that may have complex rules and regulations they need to follow. In order for aid workers and NGOs to be able to administer financial literacy training, they must themselves be aware of the complexities of financial aid.

Gender disparities within, for example phone ownership as well as digital and financial literacy were raised as well. Particularly factors such as responsible use of money and using the assistance for its intended purpose were credited to women. This is an issue that requires more empirical research and data, so that appropriate programs can be set up. Finally, Uganda has an effective registration system for refugees, and this was raised in conjunction with the discussion of digital and any type of IDs for refugees. This is also an issue between refugees residing in urban areas rather than camps. If they do not have an official refugee status, they may lose the right to some forms of assistance they may be entitled to.

The factors associated with the beneficiaries still need to be more investigated and validated. Therefore, this deliverable focused on the beneficiaries' issues and validated them for confirmability. In this report we validated several factors by measuring them through sub-items: the validated items **(**Financial literacy, Digital literacy, financial inclusion, Gender inclusion, Digital inclusion, Mobile ownerships SIM card ownership/accessibility, Internet connectivity, Affordability to access digital financial service and assistance, Beneficiary preference and acceptance, Ease of use for beneficiaries, Complaint redressal channel). It can be observed that the financial and digital literacy, especially for women is a significant factor that influence the effectiveness of digital cash-based assistance. Another important factor is the availability of services in the vernacular language of the refugees. This means that UIs of the digital cash-based assistance programs need to be reconfigured to match refugee's langua ge preferences. An interesting observation is that the smartphone ownership and phones with internet access is not the top priority. This can be since the internet connectivity is not as efficient as in developed nations and hence, the reliance on internet for cash transfer may not be the best option. Furthermore, smart phones are expensive, and refugees might not be ready to spend so much on a smartphone and prefer feature phones.

Practical Implication

The possibility of overcoming the difficulties of the implementation requirements must be explored according to the local context. Because the priority of "implementation requirements factors" varies according to the context, it is difficult to generalize the extent of implementation possibility based on a single specific experience. Therefore, it is crucial to conduct a thorough assessment of the local context before implementing the digital CBA program. This will help in identifying the specific contextua l factors that need to be addressed and developing strategies to overcome the challenges associated with them. The preliminary findings can help policy makers, government and humanitarian organizations when thinking of digital transition for CBA program in the context of refugee crisis in low-resource setting. However, this is very practical exception case of Uganda as the technology is booming in the region with high application of for instance mobile money in such context.

CONCLUSION

This research aims to investigate on the contextual factors for the execution of digital CBA programs in humanitarian aid contexts. The literature review and initial findings mainly focuses on the beneficiaries' side since the macro factors such as economic factors, legal factors are generally investigated and discussed in many frameworks, models and studies. The factors associated with the beneficiaries still need to be more investigated and validated. Therefore, this study focuses on beneficiaries experience. It can be observed that the financial and digital literacy, especially for women is a significant factor that influence the effectiveness of digital cash -based assistance. Another important factor is the availability of services in the vernacular language of the refugees. This means that the digital cash-based assistance programs need to be reconfigured to match refugee's langua ge

preferences the case country is multilingual and also hosted refugees from different countries. An interesting observation is that smartphone ownership and phones with internet access is not the top priority. This can be since the internet connectivity is not as efficient as in developed nations and hence, the reliance on internet for cash transfer may not be the best option. Furthermore, smart phones are expensive, and refugees might not be ready to spend so much on a smartphone and prefer feature phones.

One of the research limitations of this project that we faced was the data collection. As happens in Uganda, there are challenges with access to HOs and camp representatives. It was still a timely procedure and hectic to conduct interviews. Due to access and connection issues, some interviews have been received late, so they are not part of the analysis in this report. The study relies thus on eight interviews in addition to publicly published reports and academic literature that have been reviewed. Future research is recommended using multi-research method to complement the outcome of this paper. For instance, using a survey based on data from refugees looking into their perception of digitalization of CBA program, or longitudinal study given the specified timeframe can help organizations to view the trend of digitalization, risks, benefits, and refugee perceptions over time. Furthermore, cross-case comparison is recommended to generalize the findings of this study. Possibly, cross-country analysis can further address the differences and similarities of refugee perception given the enabling environment.

ACKNOWLEDGEMENTS

This research is supported by the Norwegian Research Council (NRC) under the project number: 325437 - FORSKER21, and project title "DigCBA: Responsible Use of Digital Cash-Based Assistance in Refugee Crises".

REFERENCES

Abdelgawad, A. A., Khan, A., & Baharmand, H. (2023). Exploring gaps in using digital delivery mechanisms for cash-based assistance in refugee crises. *International Journal of Disaster Risk Reduction* , *96*, 103907. https://doi.org/10.1016/J.IJDRR.2023.103907

Aslam, W., Ham, M., & Arif, I. (2017). Consumer behavioral intentions towards mobile payment services: An empirical analysis in Pakistan. *Market-Trziste*, *29*(2), 161–176. https://doi.org/10.22598/MT/2017.29.2.161

Burton, J. (2020). "Doing no harm" in the digital age: What the digitalization of cash means for humanitarian action. *International Review of the Red Cross*, *102*(913), 43–73. https://doi.org/10.1017/S1816383120000491

Ca LP. (2013). *Preconditions and key criteria required for effective CVA*. https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/preco nditions_and_key_criteria_required_for_effective_cva.pdf

CWG. (2020). *Cash and voucher assistance and the Covid-19 outbreak, Syria*. https://www.calpnetwork.org/wp-content/uploads/ninja -forms/2/CVA_Covid-19_Guidelines_Syria_March-2020.pdf

DG-ECHO. (2024). *Cash transfers - European Commission*. https://civil-protection-humanitarian-aid.ec.europa.eu/what/humanitarian-aid/cash-transfers_en

Doocy, S., Sirois, A., Anderson, J., Tileva, M., Biermann, E., Storey, J. D., & Burnham, G. (2011). Food security and humanitarian assistance among displaced Iraqi populations in Jordan and Syria. *Social Science and Medicine*, *72*(2), 273–282. https://doi.org/10.1016/j.socscimed.2010.10.023

ECHO. (2021). *Forced displacement: refugees, asylum-seekers and internally displaced people (IDPs)*. https://ec.europa.eu/echo/what/humanitarian-aid/forced-displacement-refugees-asylum-seekers-and-internally-displaced-people-idps_en

Heaslip, G., Kovács, G., & Haavisto, I. (2018). Innovations in humanitarian supply chains: the case of cash transfer programmes. *Production Planning and Control*, *29*(14), 1175–1190. https://doi.org/10.1080/09537287.2018.1542172

IOM. (2020). *Cash feasiblity assessment 2020; North-West Syria*. https://www.calpnetwork.org/wp-content/uploads/ninja -forms/2/IOM_CFA_external_final_compressed.pdf

Jaspars, S., & Sathyamala, C. (2021). *Digital bodies and digitalised welfare: North-South linkages in the politics of food assistance and social welfare*. https://socialprotection.org/discover/publications/digital-bodies-and-digitalised-welfare-north-south-linkages-politics-food

Jodar, J. (2021). *Transferring power to local CVA actors : Rhetoric vs Reality | The CALP Network*. https://www.calpnetwork.org/blog/transferring-power-to-local-cva-actors-rhetoric-vs-reality/

Juntunen, E. A., Kalla, C., Widera, A., & Hellingrath, B. (2023). Digitalization potentials and limitations of cash - based assistance. *International Journal of Disaster Risk Reduction* , *97*, 104005. https://doi.org/10.1016/J.IJDRR.2023.104005

Maciej Serda, Becker, F. G., Cleary, M., Team, R. M., Holtermann, H., The, D., Agenda, N., Science, P., Sk, S. K., Hinnebusch, R., Hinnebusch A, R., Rabinovich, I., Olmert, Y., Uld, D. Q. G. L. Q., Ri, W. K. H. U., Lq, V., Frxqwu, W. K. H., Zklfk, E., Edvhg, L. V,. (2020). Cash assistance: how design influences value for money. *Uniwersytet Śląski*, *7*(1), 343–354. https://doi.org/10.2/JQUERY.MIN.JS

Maghsoudi, A., Harpring, R., Piotrowicz, W. D., & Heaslip, G. (2023). Cash and Voucher Assistance along Humanitarian Supply Chains: A Literature Review and Directions for Future Research. *Disasters*, *47*(1), 42–77. https://doi.org/10.1111/DISA.12520

Maghsoudi, A., Harpring, R., Piotrowicz, W. D., & Kedziora, D. (2023). Digital technologies for cash and voucher assistance in disasters: A cross-case analysis of benefits and risks. *International Journal of Disaster Risk Reduction*, 103827. https://doi.org/10.1016/J.IJDRR.2023.103827

Mitchell, P., Bumbacher, S., & Yoshikawa, L. (2019). *Cash and Voucher Assistance and Risk in Financial Management and Compliance | Cash Learning Partnership*. https://www.calpnetwork.org/publication/cash-and-voucher-assistance-and-risk-in-fina ncial-management-

and-compliance/

Murray, S. (2016). *Can E-transfers promote inlcusion in emergencies: A case study from Ethiopia*. https://www.calpnetwork.org/wp-content/uploads/2020/01/ethiopiacasestudyfinal.pdf

Patil, P., Tamilmani, K., Rana, N. P., & Raghavan, V. (2020). Understanding consumer adoption of mobile payment in India: Extending Meta -UTAUT model with personal innovativeness, anxiety, trust, and grievance redressal. *International Journal of Information Management , 54*, 102144. https://doi.org/10.1016/J.IJINFOMGT.2020.102144

Tappis, H., & Doocy, S. (2018). The effectiveness and value for money of cash -based humanitarian assistance: a systematic review. *Journal of Development Effectiveness*, *10*(1), 121–144. https://doi.org/10.1080/19439342.2017.1363804

UNHCR. (2020). *MENA Quarterly Regional Cash Assistance Monitoring Update April to June 2020 (Q2)* . https://reliefweb.int/report/world/unhcr-mena-quarterly-regional-cash-assistance-monitoring-update-april-june-2020-q2

UNHCR. (2023). *UNHCR - Refugee Statistics*. https://www.unhcr.org/refugee-statistics/



Analytical Modeling and Simulation

Track Chairs

Christopher Zobel, Virginia Tech

Andrew Arnette, Virginia Tech

Josey Chacko, Mount St. Mary's University

Duygu Pamukcu, HEC Montreal

Possibility of Suppressing Fire Spread in the Event of a Large Earthquake by Greening and Demolishing Vacant Houses

Toshihiro Osaragi
Tokyo Institute of Technology
osaragi.t.aa@m.titech.ac.jp

Asuka Harada
Tokyo Institute of Technology
harada.a.aa@m.titech.ac.jp

Takuya Oki
Tokyo Institute of Technology
oki.t.ab@m.titech.ac.jp

Yoho Sakamoto
Tokyo Institute of Technology
sakamoto.y.al@m.titech.ac.jp

ABSTRACT

In recent years, disaster mitigation measures by greening and demolition of vacant houses to prevent large post-earthquake fire spread have been promoted. However, few studies evaluate its effects on reducing property damage and human casualties, especially in vulnerable areas for earthquakes, such as densely built-up wooden residential areas. In this paper, we develop a model to describe the delay of post-earthquake fire spread between neighboring buildings, and incorporate it into the simulation model. Using the integrated model, we evaluate the reduction of the number of burnt-down buildings and human casualties in densely built-up wooden residential areas, and demonstrate the effectiveness of greening and demolition of vacant houses.

Keywords

greening, demolition of vacant houses, fire-spread model, large earthquake, simulation.

INTRODUCTION

Research Background

Greening policies are being promoted around the world, primarily from the perspective of preserving the living environment, with the hope of reducing CO^2 emissions and alleviating the urban heat island effect. Concurrently, with natural disasters increasing in recent years, the function of trees is being evaluated from the perspective of disaster prevention and mitigation (Shin et al., 2014). The role of trees as windbreak forests and tidewater control forests, for example, goes without saying, but trees can also be expected to have an effect in delaying and blocking the spread of fire in urban areas. Based on a field survey of a large fire in the city of Itoigawa on December 22, 2016, Iwasaki (2019) reported that even inside the affected area were hedges with most of the leaves on the surface of the crown still green. He also reported that the houses that were spared from the spread of the fi re were adjacent to trees or empty lots. If it can be shown that greening policies are effective in suppressing the spread of fires, this could serve as an important argument for promoting greening policies in Japan. The research question of this study is validating whether urban greening can be expected to be effective in reducing fire spread damage followed by a large earthquake. Based on this background, “greening” in this paper is defined as the act of planting plants, especially evergreen or deciduous trees which have higher resistance than coniferous trees to prevent fires, between buildings to increase the effectiveness of preventing and delaying the spread of urban fires.

Recently, the problem of empty homes has worsened (Ministry of Land, Infrastructure, Transport and Tourism, 2017). The question of how to use vacant houses as housing stock is an important issue, but leaving deteriorated houses without occupants is a major problem in terms of preventing disasters and crime. Old wooden houses are

considered particularly problematic as elements that increase the risk of collapse and fire spread in the event of a large earthquake.

Because of these issues, efforts to enhance disaster prevention capabilities by greening and demolishing vacant houses are being made in areas such as those that are crowded with wooden housing (hereafter, densely built-up wooden housing areas). A method for quantitatively examining ways of greening and demolishing vacant houses to reduce property damage and human casualties effectively and efficiently in the event of a large earthquake is required.

Related Research

Li and Davidson (2013) extended their urban fire simulation model to include a fire ignition-spread-suppression process by taking into account firefighting activities and evaluated the effectiveness of the fire department's efforts. They also applied this model to a region in California and conducted parametric studies to investigate the key factors that influence fire spread. Khorasani and Garlock (2017) focused on the past 20 large fires following earthquakes from seven different countries, and reviewed the latest literature on post-earthquake fire mitigation techniques, as well as analytical tools for assessing fire ignition and spread. Manzello et al. (2018) suggested that research on large urban fires that occur after earthquakes and how to reduce the loss of buildings from such fires has lagged, and attempted to develop a roadmap for international research needs to reduce the risk of l arge outdoor fires. Manzello et al. (2020) reviewed experiments, models, and simulations related to firebrand generation, lofting, burning, transport, deposition, and material ignition in large-scale outdoor fire spread. They also pointed out that most of the knowledge in Japan that caused serious large-scale outdoor fires is still not available in English literature. Even if the technology or characteristics are unique to Japan, it is of great significance to communicate in English the experience and knowledge cultivated in Japan, which is prone to earthquake disasters. Vitorino et al. (2024) presented a review of the state-of-the-art on previous post-earthquake fire events, its risk estimation, fire ignition models, and probabilistic loss assessment. They also introduced several mitigation strategies that can help reduce or prevent post-earthquake fires.

Large-scale urban fires have occurred not only in Japan but also in various parts of the world, leading to the development of numerous sophisticated fire spread models. However, Japan's unique characteristics, including building structures (degree of fireproofing), open space sizes, road widths, distances between buildings, and the density of vegetation between structures, pose challenges in directly applying these validated models. Additionally, the rarity of large-scale fire outbreaks necessitates estimating parameters for current models based on surveys such as those conducted for the Great Hanshin-Awaji Earthquake (1995). Consequently, while we will be incorporating some well-established models rooted in earlier research, our focus will remain on post-earthquake fires occurring within Japan's distinctive environment.

The Ministry of Construction’s Comprehensive Technology Development Project (Technology Research Center for National Land Development, 1983) proposed a pioneering method of determining the time required for fire spread, the fire-blocking effect of structures, etc., by calculating the amount of radiant heat when a fire spread between two buildings by using a theoretical equation (solid angle projection ratio). Based on this method and the results of surveys of disasters such as the Great Hanshin-Awaji Earthquake, a model (Tokyo Fire Dept. Model 2001) for describing the spread of fire in urban areas more simply has been proposed (Tokyo Fire Department, 2001).

Excellent research on the fire-preventive function of trees during large fires in urban areas has been performed by Saito and Iwakawa (1982) and Iwakawa (1984). Using experiments and a review of the literature, their research has clarified the relationship between tree shape and heat shielding rate, heat shielding ability, ignition/burning conditions, and fire resistance threshold. Also, Saito and Tabata (1991) incorporated Iwakawa’s approach for calculating the shielding rate of trees into a method for calculating radiant heat in the Ministry of Construction’s Comprehensive Technology Development Project (Technology Research Center for National Land Development, 1983), and examined theoretically the fire-blocking effect as a function of tree height, crown size, etc.

Based on these theoretical models, simulations of fire spread in urban areas that take into account the presence of trees have recently been attempted. For example, using an earthquake fire spread simulation system that incorporates the fire prevention effect of trees, Futagami et al. (Futagami et al., 2006; Futagami and Kunikata, 2015; Futagami, 2023) quantitatively evaluated the fire prevention effect of trees (differences in the presence or absence of trees) by varying wind speed, wind direction, and fire point. Also, Nishi and Futagami (2019) have developed a data input interface for the earthquake fire spread simulation system that takes into account tree species and shapes such as evergreens, deciduous trees, and hedges, and used this to evaluate the effects of existing greenery, hedge placement, setbacks, the removal of vacant houses, tree planting, and planting trees around open spaces. Nagano et al. (2008) targeted densely built-up wooden housing areas in Kanazawa City to examine the

effect of tree belts on preventing fire spread. Their simulation assumed the current state of planting and additional new planting. This was similar to the simulation attempted in the present study and served as a reference. Kumagai et al. (2009) simulated fire spread in an urban area using detailed vegetation distribution data collected by remote sensing.

In research on the effect of demolishing vacant houses on disaster mitigation, Nakai et al. (2012) calculated and compared arrival rates at temporary evacuation centers in densely built-up areas of Osaka City when the sites of vacant houses are converted into roads, used for road widening, or used as temporary evacuation centers. Nakajima et al. (2018) developed a model for predicting the occurrence of vacant houses on the basis of a field survey of an area where fireproofing was being promoted in Yokohama City and made predictions for up to 50 years in the future. They used a fire-spread cluster (Kato et al., 2006) approach to estimate the reduction in the rate of buildings destroyed by fire in the event of a large earthquake that can be expected when vacant wooden houses have been fireproofed and when all vacant houses have been demolished.

As reflected in the abovementioned prior research, many theoretical studies about the effects of trees on delaying and blocking fire spread have been carried out, as well as many simulations of the effects of greening and demolishing vacant houses on reducing fire damage. However, as far as the authors of the present paper are aware, there has been no research that examined the contribution of greening and demolishing vacant houses to reducing human casualties by linking fire damage and the evacuation behavior of residents. We were also unable to find studies that quantitatively examined effective and efficient methods for determining tree placement and vacant house demolition.

Research Objectives

Osaragi and Oki (2017) constructed a comprehensive simulation model that takes into account the interaction between property damage and human behavior (evacuation behavior, rescue activities, and initial firefighting activities) during a large earthquake (hereinafter, property damage/human behavior simulation model). A feature of this model is the incorporation of initial firefighting by local residents. In densely built-up wooden housing areas, the time it takes for a fire to spread from the building of fire origin to an adjacent building (fire-spread time) is short, and initial firefighting by local residents is less likely to succeed (Oki and Osaragi, 2018). If fire spread can be delayed by greening and demolishing vacant houses, it may be possible to improve the success rate of initial firefighting and reduce the number of fatalities and fire-damaged buildings. This was the hypothesis of the present study.

In this paper, we examine a method of incorporating the effect of delaying the time of fire spread between adjacent buildings by greening and demolishing vacant houses into the property damage/human behavior simulation model. Next, we examine the effect that delaying the time of fire spread between adjacent buildings has on reducing fire damage (number of fire-damaged buildings) and human casualties (number of fatalities) in the event of a large earthquake. Finally, we examine effective and efficient tree placement and vacant house demolition.

SIMULATION MODEL FOR EVALUATING THE EFFECT OF TREES ON DELAYING FIRE SPREAD

Method of Calculating Fire-spread Delay due to Trees

Methods proposed by the Ministry of Construction's Comprehensive Technology Development Project (Technology Research Center for National Land Development, 1983), Saito and Iwakawa (1982) and Iwakawa (1984) calculate the fire-spread time when there are trees between the building of fire origin and an adjacent building (hereafter, "Iwakawa model"). Table 1 and Fig. 1 show the variables used in the Iwakawa model. With these variables, it is possible to calculate the amount of radiant heat received between the nearest adjacent poi nts of the building of fire origin and an adjacent building, R [kcal/(m^2·h)], the amount of heat received at a heated point, q [W/cm^2], and the time required for ignition, t [s] (Equations (1)–(3)). Iwakawa (1984) describes the derivation methods in detail, and because these are key equations in this paper, APPENDIX provides a summary. In Equation (1), E is the amount of radiant heat from the radiation test surface (= 44,000 kcal/(m ^{2}h)), Φ is an urban area coefficient described by building area ratio m, F is the configuration factor for the heated point of the radiation test surface that is visible through the trees, e is the mean bright surface ratio of the radiation test surface corresponding to the configuration factor $f(a, b, r)$. Also, transmittance τ is a value determined according to tree type, tree shape, and season.

Table 1. List of variables used in the Iwakawa model

	Variables
E	Radiation from the surface of flame (constantly 44,000) [kcal/(m^2· h)]
h_1	Height of tree [m]
B	Width of burning building [m]
c	Height of heated point [m]
U	Wind speed [m/s]
α	Wind direction [rad]
d	Distance between two buildings [m]
d_0	Distance between heated point and tree [m]
D	Simultaneous burning area [m]
c_1	Ratio of sum of fire-resistant buildings area divided by sum of all buildings area
c_2	Ratio of sum of semi-fire-resistant buildings area divided by sum of all buildings area
m	Building coverage ratio in the town
τ	Transmittance (= 1 - (shielding rate by a tree or a hedge))
R	Amount of radiant heat received between the nearest adjacent points of the building of fire origin and an adjacent building [kcal/(m^2·h)]
q	Amount of heat received at a heated point [W/cm^2]
t	Time required for ignition [s]
Φ	Urban area coefficient described by building area ratio m
F	Configuration factor for the heated point of the radiation test surface
e	Mean bright surface ratio of radiation test surface corresponding to configuration factor
τ	Value determined according to tree type, tree shape, and season
I, J	Heating-side building, Heated-side building
i, j	Heating-side building point, Heated-side building point
N_{IJ}	Set of surface combination
n_{IJ}	The number of elements in the set N_{IJ}
F_{IJ}	Value of the configuration factor for heating-side building I and heated-side building J
F_{ij}	Value of the configuration factor for heating-side point i and heated-side point j
R_{ij}	Amount of heat received by surface j from surface i
R_{Ij}	Amount of heat received by surface j from heating-side building I
R_{IJ}	Maximum value of R_{Ij} over j
t_{IJ}	Fire-spread time from building I to building J
D_0	Simultaneous burning area [m]
h_0	Flame height [m]
θ	Flame tilt [rad]

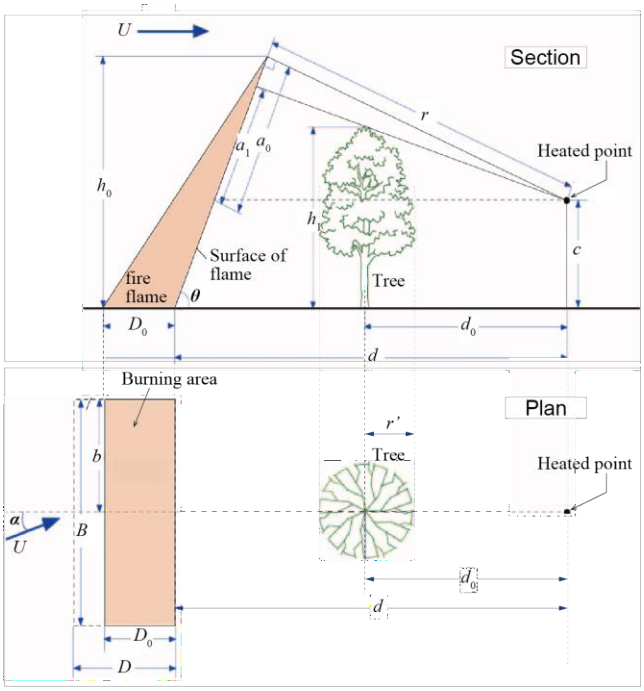


Figure 1. Illustration of variables used in the Iwakawa model

$$R = E \times \Phi^{0.5} \times F \times e \times \tau \quad [\text{kcal/(m}^2\text{h)}] \tag{1}$$

$$q = R \times 1.163 \times 10^{-4} \quad [\text{W/cm}^2] \tag{2}$$

$$t = -1.6 \times 10^3 \times \log_{10}\left(1 - \frac{0.53}{q}\right) \quad [\text{S}] \tag{3}$$

In this study, a more sophisticated model was created by incorporating consideration of the buildings on the heating-side and heated-side as face elements when calculating the amount of radiant heat received, *R*, rather than considering only the possibility of fire spread between the nearest adjacent points of adjacent buildings (Fig. 2).

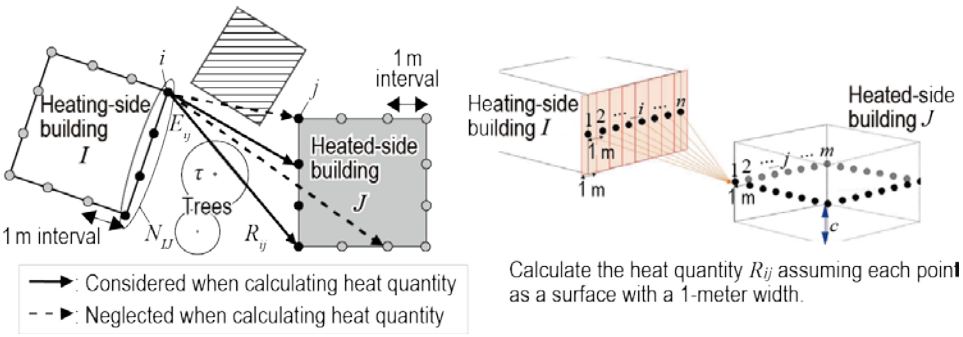


Figure 2. How to consider heating surface and heated surface for calculating the amount of radiant heat received

Specifically, we consider a heating-side building *I* and a heated-side building *J*. First, surfaces *i* and surfaces *j* are formed by dividing the outer circumference of each building into approximately 1 m intervals. Points are generated at equal intervals on the outline of the building using the "Densify" tool of ArcGIS. The interval between points is automatically adjusted to be proximate to the specified interval (here, 1 m) such that the start and end points of the point generation coincide. It is assumed that every surface *i* can be a radiating surface and every surface *j* can be a heated surface. The set of surfaces *i* is denoted by N_{IJ} and the number of elements in the set is denoted by

n_{IJ}.

CoRe Paper – Analytical Modeling and Simulation
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

If the line segment *i–j* established between surface *i* and surface *j* intersects with a building, etc., radiant heat along *i–j* is considered to be blocked and so is not included in the thermal calculations. In this case, the configuration factor for heat received from the area of the radiation test surface *i* ($\in N_{IJ}$) (Fig. 2, right), F_{ij}, can be expressed as follows:

$$F_{ij} = \frac{F_{IJ}}{n_{IJ}} \tag{4}$$

where F_{ij} is the value of the configuration factor, *F*, for heating-side building *I* and heated-side building *J*. In this case, from Equations (1) and (4), the amount of heat received by surface *j* from surface *i*, R_{ij}, is

$$R_{ij} = E \times \Phi^{0.5} \times F_{ij} \times e \times \tau \quad [\text{kcal}/(\text{m}^2\text{h})] \tag{5}$$

When the distance from the center of a tree to the path of fire spread (line segment *i–j*) is shorter than the radius of the tree, the shielding effect, $\tau < 1$, of the tree is considered. Next, the sum of R_{ij} giving the amount of heat received by surface *j* from heating-side building *I* is found, denoted R_{Ij} [kcal/(m²·h)]), and the maximum value of R_{Ij} over *j*, denoted R_{IJ}, is taken to be the amount of heat that heated-side building *J* receives from heating-side building *I* (Equation (6)). This is used to calculate the fire-spread time from building *I* to building *J*, t_{IJ} [s] (Equations (7) and (8)).

$$R_{IJ} = \max_j R_{Ij} = \max_j \sum_{i \in N_{IJ}} R_{ij} \quad [\text{kcal}/(\text{m}^2\text{h})] \tag{6}$$

$$q_{IJ} = R_{IJ} \times 1.163 \times 10^{-4} \quad [\text{W}/\text{cm}^2] \tag{7}$$

$$t_{IJ} = -1.6 \times 10^3 \times \log_{10}\left(1 - \frac{0.53}{q_{IJ}}\right) \quad [\text{s}] \tag{8}$$

Using the above method, the fire-spread time between adjacent buildings taking into account the presence of trees is calculated, and the difference in fire-spread time, Δt, is regarded as the fire-spread delay attributable to trees.

Generally, as a fire spread, heat is received from multiple burning buildings, and therefore the fire-spread time becomes shorter. However, this paper discusses the initial stage of fire spread from the building where the fire breaks out to an adjacent building, not fire spread between multiple buildings.

Incorporation into Urban Fire-spread Model

Using the model developed by Tokyo Fire Department's Fire Prevention Council (Tokyo Fire Dept. Model 2001), this paper describes the spread of fire not only between adjacent buildings but also at the urban area level. In Tokyo Fire Dept. Model 2001, fire-spread time is calculated using parameters that include the distance between adjacent buildings, wind speed, and building structure on heating-side/heated-side. Fire-spread time for between adjacent buildings can be calculated without performing complicated thermal calculations, but it is not possible to directly consider changes in the amount of heat received due to the shielding effect of trees (delay of fire spread).

Therefore, by finding the difference in fire-spread time between with and without trees, Δt, with the method described in the preceding section and adding this value to the fire-spread time, *T*, calculated using Tokyo Fire Dept. Model 2001, the shielding effect of trees can be incorporated into the fire-spread model at the urban area level. Here, two parameters (height of heated point and opening ratio) were adjusted so that the fire-spread time between adjacent buildings without trees approximated the calculation results of Tokyo Fire Dept. Model 2001 (Fig. 3). For wooden buildings, the difference between the two models is minimized by taking the height of the heated point to be half the height of the building on the fire origin side (3 m above ground). For fire-resistant and quasi-fire-resistant buildings, the height of the heated point is set at 3 m above the ground and the opening ratio (percentage of wall height taken up by opening height) at 55%. At a distance between adjacent buildings of l ess than 1 m, the difference from Tokyo Fire Dept. Model 2001 becomes rather large, but this was not considered to be an issue within the realistic range of distances between adjacent houses (Fig. 3). The rationale behind setting specific parameters for particular building combinations stems from the fact that the likelihood of fire spreading is influenced by the combination of building structures on the heating and heated sides. Specifically, wooden buildings have a higher likelihood of catching fire from adjacent burning structures compared to fire-proof buildings. Additionally, the amount of heat generated during combustion, which contributes to the spread of fire to neighboring buildings, is greater for wooden buildings and lesser for fire-resistant buildings.

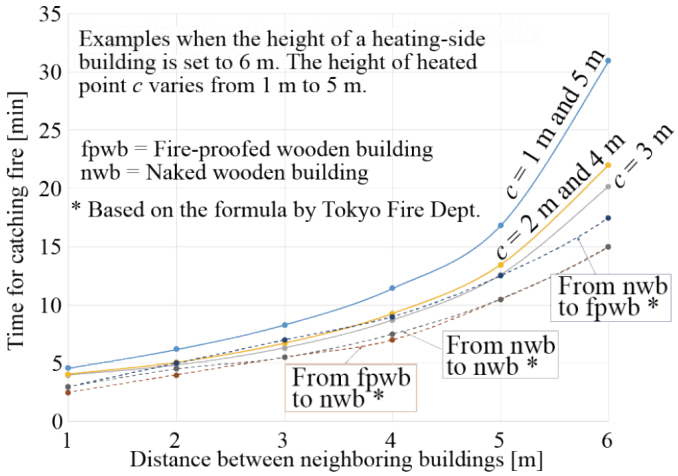


Figure 3. Distance between neighboring buildings and time for catching fire (in case of fire spread between wooden buildings)

Also, when fire spread enlarges, flame height h_0 exceeds 10 m and flame tilt θ also increases due to the effect of wind speed U and wind direction α, so a shielding effect cannot be expected from groups of trees such as those considered in this paper. Therefore, the effect of trees is not taken into account in calculations concerning the stage of fire-spread enlargement that follows the initial stage.

Method of Estimating the Number of Fire-damaged Buildings and Fatalities

The abovementioned effect of trees in delaying or blocking fire spread is incorporated into the property damage/human behavior simulation model (Tokyo Fire Department, 2001) used to examine property damage (number of fire-damaged buildings) and human casualties (number of fatalities) in the event of a large earthquake. This property damage/human behavior simulation model is a major part of the calculation process appearing in this paper. Osaragi and Oki (2017) developed a comprehensive simulation model (Fig. 4) that integrates property damages (building collapse, fire spread, and street blockage) and various activities (rescue activity, firefighting activity, and wide-area evacuation activity) by residents in a large earthquake. This model can analyze the effect and risk of rescue and firefighting activities carried out by residents under the assumption that a large earthquake directly hits the Tokyo area. Furthermore, it can analyze situations where there are many fatal casualties caused by people failing to evacuate, being trapped on streets, and so on. The details of the model is given in Osaragi and Oki (2017).

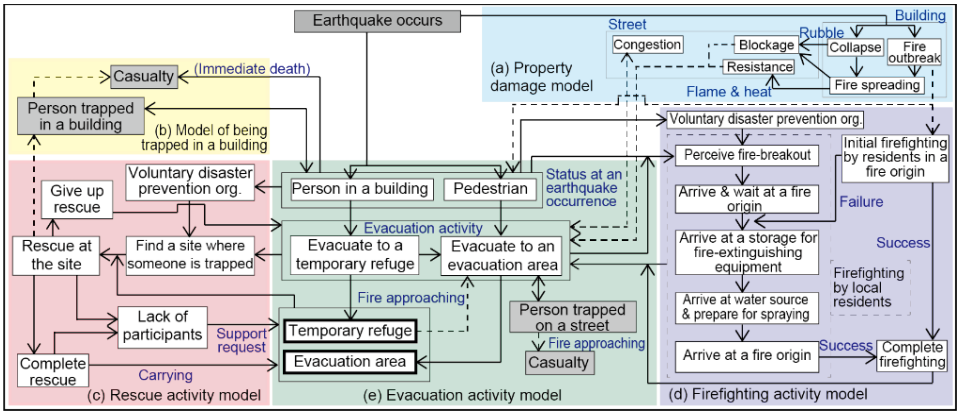


Figure 4. Overview of the property damage/human behavior simulation model (Osaragi and Oki, 2017). This is a comprehensive simulation model that integrates property damages (building collapse, fire spread, and street blockage) and various activities.

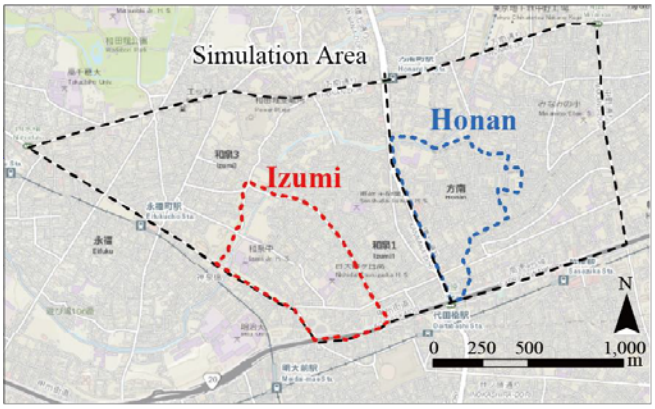
However, many pages would be required for a detailed explanation, and therefore a summary of the model is given in Osaragi and Oki (2017).

The following sections analyze the extent to which the number of fire-damaged buildings and fatalities due to a large earthquake could be reduced by reducing the fire-spread time through greening and demolishing vacant houses. Specifically, it is expected that by reducing the fire-spread time, initial firefighting by local residents using standpipes, for example, could succeed in time to limit the fire damage to the building of fire origin only. A reduction in the number of fire-damaged buildings could reduce the risk of people becoming trapped in buildings or streets being engulfed by the fire. In the Kobe Earthquake (1995), approximately 60% of the post-earthquake fires were attributed to electrical faults, occurring within 4 days following the seismic event. These fires erupted after residents had evacuated. In response to this challenge, the Japanese government has been advocating for the implementation of seismic breakers, which automatically cut off power circuits upon detecting seismic activity. This research focuses on other fires (40% of post-earthquake fires) that emerge immediately after an earthquake and are identified by nearby residents or locals before/during evacuation. When a large earthquake strikes, water pipes may be damaged, and the impact of this cannot be ignored (Osaragi et al., 2022). However, to avoid the postulated event becoming too complicated, this was not considered in the present study.

EFFECT OF EXISTING TREES ON REDUCING FIRE DAMAGE

Analysis Areas and Simulation Scope

Honan 1-chome, Suginami-ku (Honan area), which is one of the special fireproof zones in Tokyo's "Ten-Year Project to Advance Fire Resistance in Densely Built-up Wooden Housing Areas" (Tokyo Metropolitan Government, 2023a) and is also considered to be at high risk in the "8th Community Earthquake Risk Assessment Study" (Tokyo Metropolitan Government, 2023b), and nearby Izumi 2-chome, Suginami-ku (Izumi area), which has a relatively large amount of trees, were chosen as the analysis areas (Fig. 5). However, because fire could also spread from outside of these areas, an area surrounded by wide streets that could act as firebreak belts was taken as the simulation scope (Fig. 5). Evacuees were considered to have completed their evacuation when they arrived at any one of these wide streets.



	Izumi area
Number of buildings	1,473
	Honan area
	1,769
Number of people	Izumi + Honan
	6,684

People in buildings and outdoor pedestrians only in Honan and Izumi areas at 6:00 pm on weekdays are considered.
The number of people is estimated by using the Person Trip Survey Data in TMA (2008). The initial position of person in buildings and outdoor pedestrians are estimated by each building and street respectively 20),21).

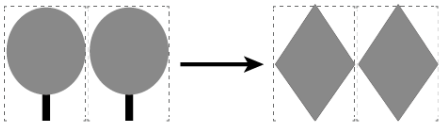
* Although property damages are considered in the whole simulation area, fire-outbreaks and trees are considered only in Honan and Izumi areas.
* When a person arrives at any wide streets shown by black broken line, he/she is considered to complete evacuation.

Figure 5. Study area: one of the special fireproof zones in Tokyo's "Ten-Year Project to Advance Fire Resistance in Densely Built-up Wooden Housing Areas"

Simulation Assumptions

Table 2 shows the simulation assumptions. The "current state" was taken as 2011 for buildings, while for trees, it was taken as the time of implementation of the field survey (2018). Data from 2011 building surveys (ward areas) provided by the Tokyo Metropolitan Government's Bureau of Urban Development were used here. The height of existing trees (three classes: 2 m, 4 m, and 6 m) and the type of tree (evergreen or deciduous) were judged by eye from the field survey (locations that were difficult to see were supplemented by aerial photographs). Tree width was assumed to be half of tree height. Figure 11 (discussed in detail later) shows the spatial distribution of existing trees and buildings (according to fire preventive construction).

Table 2. Assumptions needed for performing simulation analysis

PGV (Peak Ground Velocity)	66 cm/s
Seismic intensity	5.96
Season and weather	Sunny day in winter
Earthquake occurrence time	6:00 pm on a weekday
Wind speed and wind direction	3 m/s, north wind
Shielding rate by a hedge	(Evergreen tree) 0.7837, (Deciduous tree) 0.2105
Shielding rate by a tree (except a hedge) *	(Evergreen tree) 0.7837 x 0.5 = 0.3919 (Deciduous tree) 0.2105 x 0.5 = 0.1053
	* Regarding the part covered by a tree (except a hedge) as a rhombus as shown in the left figure, it can be treated as a barrier wall with half of the area obtained by multiplying the height and width of the tree. Therefore, the shielding rate of a tree trunk multiplied by 0.5 is used.

Effect of Existing Trees on Reducing Number of Fire-damaged Buildings/Fatalities

There are 41,281 building pairs with the potential for fire spread inside the analysis areas (Fig. 6). Because the property damage/human behavior simulation requires a long computation time, the following procedure was first used to narrow down the buildings to those with the potential for reduced damage, in order to efficiently verify the effect of trees on reducing damage. (1) The delay of fire spread between adjacent buildings was calculated for those with trees only. (2) The effect of delaying or blocking fire spread was confirmed in 226 pairs. (3) These 226 pairs were configured by taking 151 buildings as heating-side buildings. (4) Considering there to be a potential for damage reduction when a fire starts from these 151 buildings, fire-spread simulations were performed taking each of the buildings as the building of fire origin. An effect was found for 10 of these 151 buildings (when the fire spreads more quickly between adjacent buildings without trees, the trees do not work to delay the fire spread). (5) Assuming that a fire starts from one of these 10 buildings, the property damage/human behavior simulation was run until 24 hours after the outbreak for 100 cases with different roadblock situations (10 buildings × 100 cases = 1,000 simulations in total).

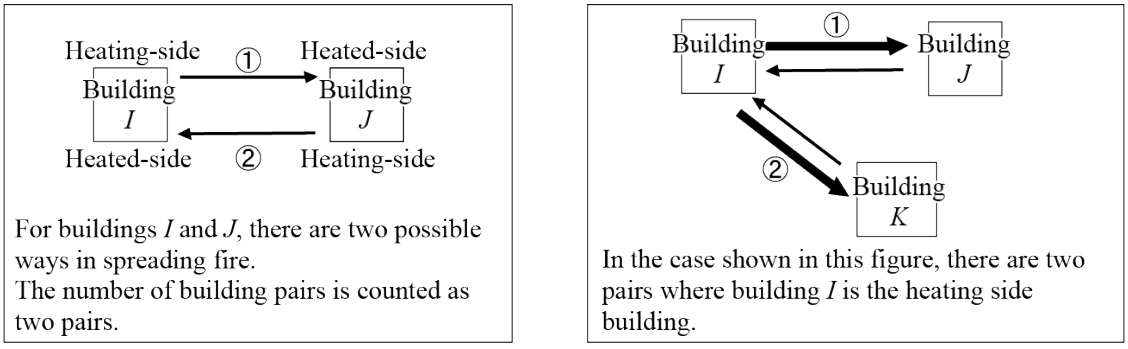
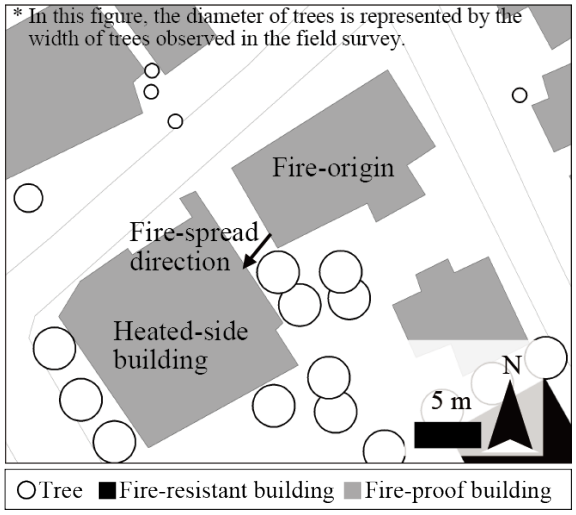


Figure 6. Definition of building pairs with the potential for fire spread inside the analysis areas

Two of these 10 buildings were cases in which firefighting activities succeeded due to the trees delaying the fire spread. Large disaster mitigation effects were confirmed in cases where the roadblocks around the building of fire origin and equipment stores were minor and rapid initial firefighting was possible. Table 3 shows the specific results. For the building of fire origin 1, the fire-spread time was delayed by the trees from 19.1 minutes to 20.8 minutes (Fig. 7). This difference of 1.7 minutes is not large, but initial firefighting activities were successful in 17 of 100 cases, and as a result, the number of fire-damaged buildings was reduced from 2,003 buildings to 1 building (the building of fire origin only). Similarly, for the building of fire origin 2, initial firefighting activities were successful in 26 cases, and along with a substantial decrease in the number of fire-damaged buildings, the average number of fatalities was successfully reduced by approximately 13, and the maximum number was reduced by 58.

Table 3. The numbers of burnt-down buildings and casualties with and without the present trees

* Circled numbers ① and ② indicate the location of each fire-origin building.		Result of fire-extinguishment	①	②
Number of successful fire-extinguishment cases thanks to trees where the fire-extinguishment are failed without trees			17/100	26/100
Number of burnt-down buildings		Failure	2,003	1,751
		Success	1	1
Number of casualties	Ave.	Failure	20.1	28.5
		Success	16.1	15.3
	Max.	Failure	44	71
		Success	21	19
Difference of the number of casualties between successful and failed fire-extinguishment cases	Ave.		-3.9	-13.2
	Max.		-24	-58



Considering the existing trees, time from the fire-outbreak to spread of fire to the heated-side building was delayed by 1.7 minutes (from 19.1 minutes to 20.8 minutes). The delay enabled firefighting activities at the building of fire-origin successful. This resulted in decreasing the number of burnt-down buildings from 2,003 to 1 (only the building of fire-origin itself).

Figure 7. Example of the place where firefighting activities were successful by considering the present trees (corresponding with Table 3①)

The success or failure of initial firefighting activities is extremely important because it greatly influences the extent of damage in densely built-up wooden housing areas. The above simulation results are interesting because they suggest that the delay of fire spread provided by trees could allow initial firefighting activities by residents to succeed. However, it is important that the accuracy of the above calculation results be carefully examined based on the various models and parameter settings incorporated into the simulation.

EFFECT OF DEMOLISHING VACANT HOUSES AND METHOD OF DETERMINING TREE PLACEMENT

Damage Reduction Effect of Demolishing Vacant Houses

This section examines the effect of delaying the spread of fire by demolishing deteriorated vacant houses to obtain empty lots. Taking Honan area as the analysis area, a total of 12 vacant houses were identified by the field survey (Fig. 8). Here too, the buildings were narrowed down using the following procedure in order to efficiently detect the effect of demolishing the vacant houses. (1) The number of fire-damaged buildings after 24 hours was calculated assuming for each of the buildings in Honan area as the building of fire origin. (2) Cases in which the number of fire-damaged buildings was reduced by demolishing the vacant houses were detected. Figure 9 shows a case in which a fire in building A spreads to a large vacant house B and ultimately causes a huge amount of damage (number of fire-damaged buildings: 1,751); however, if vacant house B is demolished, the spread of the fire can be suppressed.

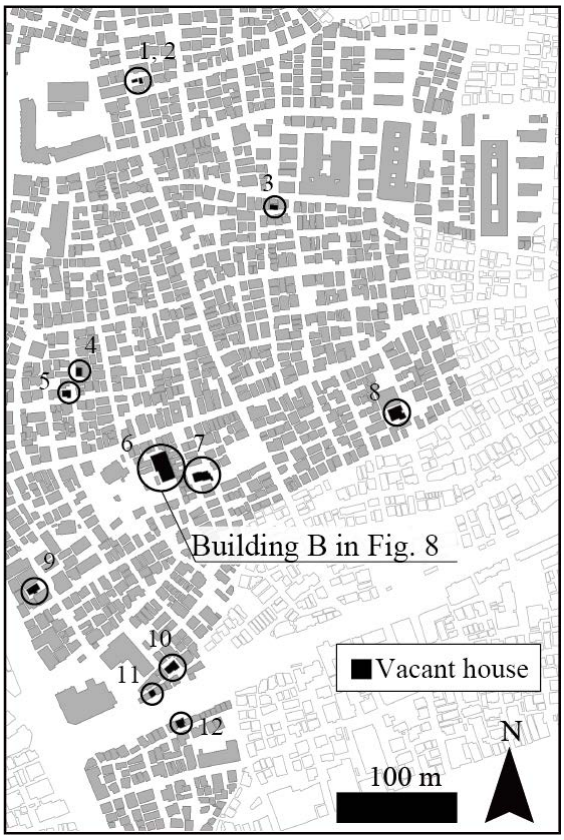


Figure 8. Spatial distribution of vacant houses in Honan area (in the present situation)

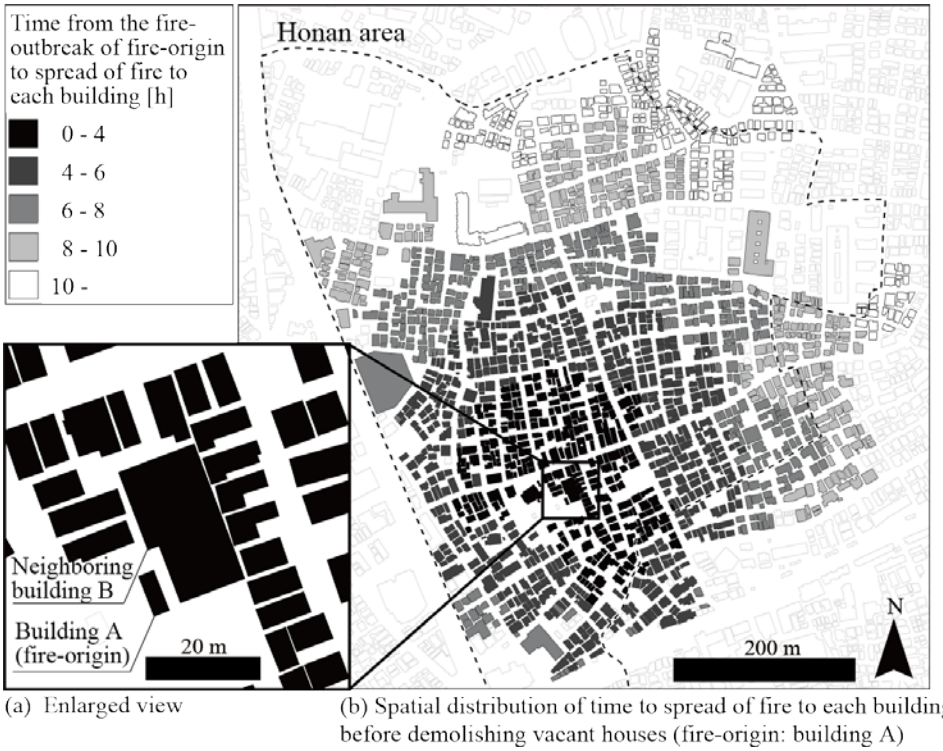


Figure 9. Example of a case where the number of burnt-down buildings is decreased due to the demolition of vacant houses

Taking the outbreak of fire in Building A as an example, the property damage/human behavior simulation was run assuming 100 cases with different spatial distributions of building collapses and roadblocks (1 building × 100 cases = 100 simulations in total). As a result, demolishing vacant house B was found to have the effect of reducing the number of fire-damaged buildings in 13 out of 100 cases (success or failure of initial firefighting was divided). Table 4 summarizes the damage situation in these 13 cases. It shows that as a result of vacant house B being demolished, human casualties in Honan area was halved due to initial firefighting by residents succeeding and preventing the spread of the fire. Simulations were conducted to assess property damage and human behavior under various scenarios involving different spatial distributions of collapsed buildings and street blockages. However, the ignition building was fixed to Building B. In the adopted fire spread model, the range of fire spread remains consistent when the ignition building remains unchanged. Consequently, in the simulated cases, only two outcomes were feasible: either the fire was confined to building A alone, or it extended widely to encompass building B. This is just a single case amid countless combinations of conditions, but it is important as an example demonstrating the importance of initial firefighting in densely built-up wooden housing areas where evacuation routes may get cut off. Building B, an aging wooden structure showing signs of deterioration, not only lacks aesthetic appeal but also poses a potential hazard as it could collapse during a major earthquake, potentially obstructing evacuation routes. Hence, its demolition has been recommended in the community. While demolishing the building incurs financial costs, the benefits of replacing it with a fireproof structure, such as a public facility or housing complex, are expected to be substantial. Moreover, given the scarcity of parks in this area, there is considerable value in transforming the space into a pocket park, providing a safe and recreational environment for the residents.

Table 4. Difference in the number of burnt-down buildings and casualties before and after the demolition of vacant houses

Number of burnt-down buildings [building]		
	Ave.	Max.
Remain	1,751	1,751
Demolish	1	1
Difference	1,750	1,750

Number of casualties [person]		
	Ave.	Max.
Remain	36.1	54
Demolish	16.5	22
Difference	19.5	39

Effect of New Tree Placement on Suppressing Fire Spread

This section examines the extent to which fire damage can be reduced by newly placing trees. Figure 10 shows the method of determining the locations of new trees.

In densely built-up wooden housing areas, it is not always possible to place trees between all adjacent buildings. For the present study, we assumed that trees can be placed if the interval between adjacent buildings is at least 2.5 m (Table 5). Considering placement within residential premises, with reference to Article 128 of the Order for Enforcement of the Building Standards Act (pathways within premises), it was considered necessary to leave a distance of at least 1.5 m from one of a pair of adjacent buildings. Also, based on “greening using medium-sized trees in front of walls” in the Guidebook for Planting Greenery (Setagaya Ward, 2021), the distance from the other building was taken to be 0.5 m. Further, considering the minimum tree width to be 0.5 m (hedge), the interval between adjacent buildings in which it is possible to place trees was taken as 2.5 m or greater (Table 5). However, it is not practical or efficient to place trees wherever placement is possible. Therefore, the following method was used to narrow down the placement locations. Calculate the fire-spread delay between all adjacent buildings where tree placement is possible and a damage reduction effect can be expected (Fig. 10, Step A). Establish an order of priority based on the length of fire-spread delay (Fig. 10, Step B). In order starting from the place with the highest priority, place a certain number of trees using the method shown in Fig. 10, Step C. Basically, trees are placed according to the following rules. Rule 1: Do not place trees where they would overlap with streets or pathways. Rule 2: Plant trees that are as large as possible (Table 5). Rule 3: Plant the minimum number of trees required to block the path of fire spread. Rule 4: When it is possible to place trees on either the heated-side or the heating-side, place them on the side that is most effective.

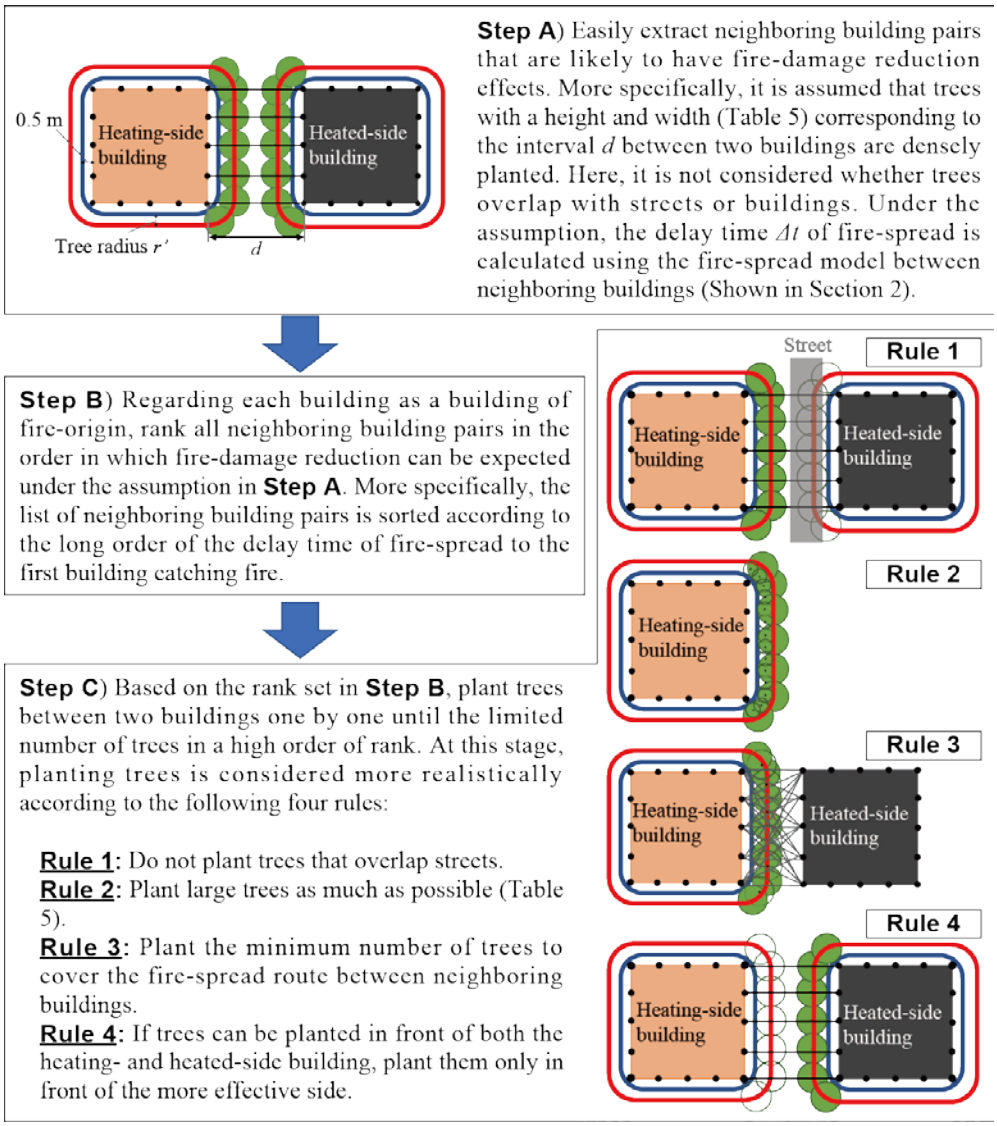


Figure 10. How to plant new trees: methods of determining the locations of new trees

Table 5. Assumption for the interval between two buildings, tree height, and tree width

	Interval d between two buildings	Tree height h_1	Tree width ($2r'$)
Level 0	Less than 2.5 m	Impossible to plant trees	
Level 1	2.5 m or more but less than 4 m	2 m	0.5 m (hedge)
Level 2	4 m or more but less than 5 m	4 m	2 m
Level 3	5 m or more	6 m	3 m

Based on the above assumptions, trees equivalent to approximately 10% of the existing trees (total 894) were placed, and (i) a method in which the trees are placed at random between adjacent buildings and (ii) the method described above in which trees are placed according to order of priority were compared. Specifically, 100 cases with different distributions of building collapses and roadblocks were analyzed (Fig. 11). In assumption (i-1), a total of 96 trees/hedges were placed between 12 pairs of adjacent buildings; in assumption (i-2), a total of 92 trees/hedges were placed between 10 pairs of adjacent buildings; and in assumption (ii), a total of 93 trees/hedges were placed between 8 pairs of adjacent buildings.

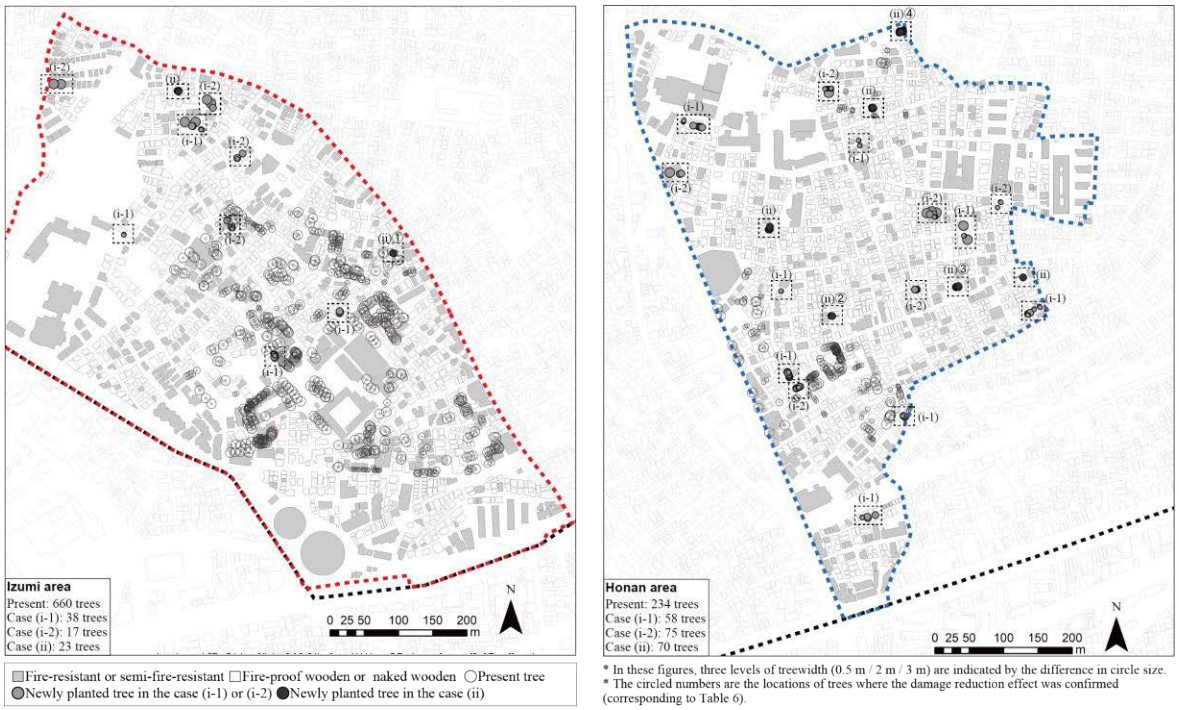


Figure 11. Spatial distribution of present/newly placing trees and the spatial distribution of existing buildings (according to fire preventive construction)

In placement method (i), fire-spread time to an adjacent building was delayed only slightly, and cases in which initial firefighting succeeded due to the addition of trees could not be confirmed. However, in placement method (ii), the effect of the trees was confirmed when the building of fire origin was one of a certain four buildings. Table 6 shows the numbers of fire-damaged buildings and fatalities in these cases (total of 19 cases). The numbers vary depending on the building collapse and roadblock situation, but a case in which the number of fatalities decreased by 89 due to the placement of trees was confirmed (Table 6, Location 3).

Table 6. The number of burnt-down buildings and casualties before and after planting new trees

* Circled numbers ① to ④ indicate the location of each fire-origin building (Fig. 12).		Planting new trees	①	②	③	④
Number of successful fire-extinguishment cases thanks to new trees where the fire-extinguishment are failed before planting new trees			3/100	12/100	2/100	2/100
Number of burnt-down buildings		Before	2,241	1,751	1,751	1,751
		After	1	1		1
Number of casualties	Ave.	Before	19.0	30.1	63.5	29.5
		After	16.7	18.7	15.5	22.0
	Max.	Before	25	52	102	36
		After	19	30	18	22
Difference of the number of casualties between before and after planting new trees	Ave.		2.3	11.4	48.0	7.5
	Max.		6	35	89	14

SUMMARY AND CONCLUSIONS

In this paper, we examined the possibility of reducing damage by delaying the spread of fire between adjacent buildings through greening and demolishing vacant houses, thereby improving the success rate of initial firefighting by local residents. We also examined methods of determining tree placement to reduce damage efficiently. The main results are given below.

First, after summarizing a previous model of fire spread between adjacent buildings and a procedure for calculating the amount of heat received taking into account trees, we developed a more sophisticated model of fire spread between adjacent buildings that assumes not only fire spread between the nearest adjacent points of the heating-side building and the heated-side building, but also fire spread between the heating surface and heated surface. Also, by examining a method of incorporating this model into Tokyo Fire Dept. Model 2001, which is a proven simulation model for fire spread in urban areas, we made it possible to quantitatively evaluate the effect of trees on delaying fire spread. Furthermore, through integration with a simulation model for property damage/human behavior in the event of a large earthquake, we made it possible to evaluate the effect of trees on reducing the number of fire-damaged buildings and fatalities due to a large earthquake.

Using the property damage/human behavior simulation model, we analyzed densely built-up wooden housing areas in Tokyo and evaluated the effect of the existing trees on reducing damage. The results confirmed that, in some cases, trees delay the spread of fire between adjacent buildings and allow initial firefighting by local residents to succeed, leading to a reduction in the number of fire-damaged buildings and fatalities. We also examined the disaster mitigation effect of demolishing vacant houses and confirmed that, in some cases, the number of fatalities could be halved by demolishing vacant houses that create a path of fire spread from the building where the fire breaks out.

Another characteristic of densely built-up wooden housing areas is that, although it is unlikely, if initial firefighting does succeed, an extremely large disaster mitigation effect is obtained. Using the simulation system developed in this paper, it is possible to discover where in an analysis area it would be efficient to newly place trees. We consider this to be a valid method for making effective use of limited space and budgets to improve the safety of an area and simultaneously enrich the green environment of people's everyday lives.

However, these disaster mitigation effects were observed under limited conditions, and effects on the order of our original expectations were unfortunately not obtained. This is considered to be due to the particular spatial characteristics of densely built-up wooden housing areas. Specifically, in densely built-up wooden housing areas, because the distances between adjacent houses are extremely small, the speed of fire spread (fire-spread time) is determined not by adjacent houses with trees but by adjacent buildings with small gaps between them. Trees do not originally exist in these small gaps, and it is difficult to newly place trees in them. Also, in densely built -up wooden residential areas, there are many structurally weak houses with low earthquake resistance located along narrow streets; consequently, many roads become blocked in the event of a large earthquake. This is a factor that impedes the initial firefighting activities of local residents and is why a delay in fire-spread time due to trees cannot be exploited effectively. Comparisons with simulation analyses targeting areas other than densely built-up wooden residential areas need to be performed to examine the possibilities of the disaster mitigation effects of trees and vacant house demolition. Hence, the experimental results obtained in this research are limited by a narrow set of conditions. The authors acknowledge these limitations and would like to incorporate sensitivity analyses to assess the robustness of our findings under various scenarios to enhance the study's validity and strengthen the overall validity of this paper. We intend to address these in the future work.

ACKNOWLEDGMENTS

The authors would like to acknowledge the valuable comments and useful suggestions from anonymous reviewers of the Scientific Program Committee, ISCRAM 2024. This paper is part of the research outcomes funded by KAKENHI (Grant Number 23H01653).

APPENDIX

The length of the perpendicular line from the heated point to the radiation test surface r [m], urban area coefficient Φ, simultaneous burning area D_0 [m], flame height h_0 [m], and flame tilt θ [rad] are expressed by Equations (a-1)–(a-5).

$$r = \left[\left(d - \frac{c}{\tan\theta} \right)^2 - \left\{ (d + c\tan\theta)\cos\theta - \frac{c}{\sin\theta} \right\}^2 \right]^{0.5} \quad \text{(a-1)}$$

$$\Phi = 1 - 0.6(c_1 + 0.8c_2)\, m \quad \text{(a-2)}$$

$$D_0 = \min\left(D, 50, 1.179\Phi^3 + 10\right) \quad \text{(a-3)}$$

$$h_0 = 4.7 \times 0.0133 \times (D_0 - 10) \times (\Phi - 0.1) + \frac{3.98}{D_0^{0.6}} \times \left(\frac{D_0^3}{U} \right)^{0.2} \quad \text{(a-4)}$$

$$\theta = \arcsin\left(\frac{2}{U} \right)^{0.2} \quad \text{(a-5)}$$

Next, the configuration factor F for the heated point of the area of the radiation test surface is calculated. The length of the radiation test surface when there are no trees a0 [m] and the length of the radiation test surface blocked by trees a1 [m] are expressed by Equations (a-6) and (a-7).

$$a_0 = \frac{h_0}{\sin\theta} - c \quad \text{(a-6)}$$

$$a_1 = \left\{ 1 - \frac{d_0^2}{d_0^2 - (h_1 - c)^2} \right\}^{0.5} \times \frac{(a^2 + r^2)^{0.5}}{\left(\frac{d_0^2}{d^2 - (h - c)^2} \right)^{0.5} \times \sin\theta + \left(1 - \frac{d_0^2}{d^2 - (h - c)^2} \right)^{0.5} \times \cos\theta} \quad \text{(a-7)}$$

Also, the configuration factor for the surface of length *a* [m] and width *b* [m], *f*(*a*, *b*, *r*), and the mean bright surface ratio *e* of the portion of the radiation test surface corresponding to *f*(*a*, *b*, *r*) are found using Equations (a-8) and (a-9).

$$f(a,b,r) = \frac{1}{2\pi} \left[\sin\theta \left\{ \frac{a}{(a^2 + r^2)^{0.5}} \tan^{-1} \frac{b}{(a^2 + r^2)^{0.5}} + \frac{b}{(b^2 + r^2)^{0.5}} \tan^{-1} \frac{a}{(b^2 + r^2)^{0.5}} \right\} - \cos\theta \left\{ \frac{r}{(a^2 + r^2)^{0.5}} \tan^{-1} \frac{b}{(a^2 + r^2)^{0.5}} - \tan^{-1} \frac{b}{r} \right\} \right] \quad \text{(a-8)}$$

$$e = 1 - \frac{h_0 + h_1}{2h_0} \quad \text{(a-9)}$$

From Equation (a-8), the configuration factor *F* for the heated point of the area of the radiation test surface that is visible through the trees is found using Equation (a-10).

$$F = 2\left\{ f(a_0, b, r) - f(a_1, b, r) \right\} \quad \text{(a-10)}$$

REFERENCES

Futagami, T., Kimata, N. and Suehiro, F. (2006). Development of the supporting system of the fire prevention-trees maintenance based on fire-spreading risk simulator, *Infrastructure Planning Review*, 23(2), 325-334.

Futagami, T. and Kunikata, Y. (2015). Development and application research of fire thermal risk analysis system at the time of earthquake for the district disaster management plan, *Journal of Japan Society of Civil Engineers*, Ser. F6 (Safety Problem), 71(2), I_97-I_104.

Futagami, T. (2023) Systematic Study on the Fire Protection Effect of Existing Trees, J*ournal of Japan Society of Civil Engineers*, 79(24).

Iwakawa, N. (1984). Study on the fire preventive function of trees in urban areas, *Report of the Building Research Institute*, 105.

Iwasaki, T. (2019). Actual condition of damaged trees by the Itoigawa Station North Area Conflagration, *Landscape Research Japan Online*, 12, 10-13.

Khorasani, N., Garlock, M. (2017) Overview of fire following earthquake: historical events and community responses, *International Journal of Disaster Resilience in the Built Environment*, 8(02), 158-174.

Kato, T., Hong, C., Yusfu, Y., Yamaguchi, M. and Natori, A. (2006). A method for the integrated earthquake fire risk evaluation based on the single building fire probability applying to any different map-scale, *Journal of Social Safety Science*, 8, 279-288.

Kumagai, K., He, Yong and Iseki, S. (2009). Development of the spatial analysis method of vegetation distributions in an urban area in terms of prevention of the fire-spreading, *Theory and Applications of GIS*, 17(2), 45-56.

Li, S. and Davidson, R. (2013) Parametric study of urban fire spread using an urban fire simulation model with fire department suppression, *Fire Safety Journal*, 61, 217-225.

Manzello, S.L., et al., (2018) Summary of workshop large outdoor fires and the built environment, *Fire Safety Journal*, 100, 76-92.

Manzello, S.L., Suzuki, S. Gollner, M.J., Fernandez-Pello, A.C. (2020) Role of firebrand combustion in large outdoor fire spread, *Progress in Energy and Combustion Science*, 76.

Ministry of Land, Infrastructure, Transport and Tourism (2017). Akiya-tou-no-genjou-ni-tsuite (*Present conditions of vacant houses*), Dai-30-kai-fudousan-bukai-shiryou 2017, http://www.mlit.go.jp/common/001172930.pdf (Accessed 2019.10.6).

Nagano, S., Masuta, T., Hayashi, Y. and Nii, D. (2008). Study on the effect against fire-spreading of tree planning subject to large wooden houses area: Case of large wooden houses area in Kanazawa, *Proceeding of the 2008 architectural research meetings* (*Hokuriku Chapter*), Architectural Institute of Japan, 17-20.

Nakai, S., Kana, K. and Sakuma, Y. (2012). Research on actual condition of existing vacant houses and study on possibility of utilizing them as "spaces for disaster prevention" in the densely built-up area: The case of Tsuruhashi area in Osaka City, *Journal of the City Planning Institute of Japan*, 47(3), 1063-1068.

Nakajima, S., Inagaki, K., Yoshida, S. and Sadohara, S. (2018). Clearance and utilizing vacant houses for preventing fire-spreading in high density areas of wooden buildings: a case study in K-district Minami-ku Yokohama city, *Summaries of Technical Papers of Annual Meeting*, Architectural Institute of Japan, F-1. 1003-1004.

Nishi, Y. and Futagami, T. (2019) Development and Application Research of Earthquake Fire Spread Simulator Considering Green Space for Disaster Prevention Community, *Journal of Japan Society of Civil Engineers*, 75(2).

Oki, T. and Osaragi, T. (2018). Effects of firefighting activities performed by local residents to mitigate fire destruction damage and human casualties in large earthquakes, *Journal of Disaster Research* (*Special Issue on Disaster and Big Data 3*), 13(2), 272-280.

Osaragi, T. and Oki. T. (2017). Wide-Area evacuation simulation incorporating rescue and firefighting by local residents, *Journal of Disaster Research* (*Special Issue on Disaster and Big Data 2*), 12(2), 296-310.

Osaragi, T., Ogino, K., Hirokawa, A., Oki, T. (2022). Severity of Crowding at Evacuation Shelters after a Major Earthquake, *Proceedings of the 19th International Conference on Information Systems for Crisis Response and Management*.

Saito, Y. and Iwakawa, N. (1982) Fire Preventive Function of Trees: Prediction of the Effect of Fire Prevention with Trees, *Journal of the Japanese Institute of Landscape Architects*, 46(5), 158-163.

Saito, Y. and Tabata, S. (1991). A study on the effect of fire prevention by trees distributed over the urban districts of a city, *Tech. Bull. Fac. Hort. Chiba Univ.*, 44, 93-98.

Setagaya Ward (2021). Kenchiku-ni-tomonau-ryokka-no-tameno-shokusai-gaidobukku (*Guidebook for planting greenery*), https://www.city.setagaya.lg.jp/mokuji/sumai/010/004/001/d00029589.html (Accessed 2023.11.21).

Shin, Y.C., Noaki, M., Ohmiya, Y., Hayashi, Y. (2014) Experimental Study on Effect to Prevent Fire Spread Between Adjacent Buildings: −Radiant heat controlled by camellia–, *AIJ Journal of Technology and Design*, 20(44), 159-164.

Technology Research Center for National Land Development (1983). Kensetsushou-sougou-gijutsu-kaihatsu-purojekuto (*Comprehensive Technology Development Project by Ministry of Construction*).

Tokyo Fire Department (2001). Jishin-kasai-ni-kansuru-chiiki-no-bousai-seinou-hyouka-shuhou-no-kaihatsu-to-katsuyou-housaku (*Development and utilization of evaluation methods of regional disaster prevention performance for earthquake fires*), Kasai-yobou-shingikai-toushin-14 (14th report of the fire prevention council).

Tokyo Metropolitan Government (2023). Mokumitsu-chiiki-funenka-10-nen-purojekuto (*10-year project to make densely-built wooden residential areas fire-resistant*), http://www.toshiseibi.metro.tokyo.jp/bosai/mokumitu/index.html (Accessed 2023.11.21)

Tokyo Metropolitan Government (2023). Jishin-ni-kansuru-chiiki-kikendo-sokutei-chousa-dai-8-kai (*8th community earthquake risk assessment study*), http://www.toshiseibi.metro.tokyo.jp/bosai/chousa_6/home.htm (Accessed 2023.11.21).

Vitorino, H., Khiali, V., Rodrigues, H. (2024) Post-earthquake fire risk and loss assessment in urban areas, *Innovative Infrastructure Solutions*, 9(1).

Enabling Information Transmission in Low-Throughput Wireless Channels for Aerial Disaster Monitoring Systems

Julian Zobel*
Multimedia Communications Lab (KOM)
Technical University of Darmstadt, Germany
julian.zobel@tu-darmstadt.de

Ralf Steinmetz
Multimedia Communications Lab (KOM)
Technical University of Darmstadt, Germany
ralf.steinmetz@kom.tu-darmstadt.de

ABSTRACT

The utilization of Unmanned Aerial Vehicles (UAVs) for disaster relief has significantly increased in recent years, especially in the fields of aerial monitoring and communication support. At the same time, research on wireless communication channels between UAVs operating in disaster areas and their base station has been neglected, despite its importance for system operability and information transmission. Therefore, this paper addresses requirements of a suitable data channel on the use case of an Aerial Monitoring System for distributed disaster networks. We put special emphasis on possible applications using low-throughput technologies, like LoRa, to enable information transmission on low-cost hardware without centralized infrastructure, which is highly important for UAV applications in disaster scenarios. Specifically, we propose the aggregation of topology information to reduce message sizes and increase usability under challenging network conditions. Our results highlight that the aggregated monitoring information can be transferred to the base station with reasonable delay and reliability, possible with a low-throughput channel as low as 300 $\frac{bit}{s}$. At the same time, the aggregated information still maintains an adequate representation of the monitored topology of the distributed disaster network.

Keywords

Unmanned Aerial Vehicles (UAV), Aerial Disaster Monitoring, Communication Networks, LoRa, Simulations

INTRODUCTION

Disaster preparedness and disaster relief are two crucial challenges our society has faced and will always face with natural and human-made catastrophes. Especially large-scale natural disasters caused by extreme weather conditions severely increased in recent years due to climate change (Gallucci, 2018; Ranghieri & Ishiwatari, 2014; Toya & Skidmore, 2018; Universite´ Catholique de Louvain (UCL), 2022). The parallel increase of societal reliance and dependence in everyday life on sophisticated but vulnerable *Information and Communication Technologies* (ICT), like always available high-speed cellular networks, constantly aggravate these challenges during and after disasters. In recent years, several events like Hurricanes in Puerto Rico or large floods in parts of Germany highlighted that vulnerability of our existing ICT infrastructure. Severely damaged or completely destroyed landlines and cellular networks lead to a complete communication blackout. This impaired any effort for disaster relief and mitigation, especially in the first few days before provisional solutions to provide even the most basic communication capabilities, such as emergency calls, could be set up (Gallucci, 2018; German Federal Agency for Civic Education, 2021; Kreienkamp et al., 2021; Zorrilla, 2017). Since communication is vital for an efficient disaster relief, researchers all over the world emphasize the need for more robust and resilient infrastructure (Hagar, 2015; Hollick et al., 2019; Kreienkamp et al., 2021; Mori et al., 2015; Toya & Skidmore, 2018; Whittaker et al., 2013). This also includes the need for new approaches and methods to improve communication and mitigate the effects of destroyed infrastructure in a disaster situation (Fehling-Kaschek et al., 2020; Hollick et al., 2019; Mu¨hlha¨user, 2021).

*corresponding author

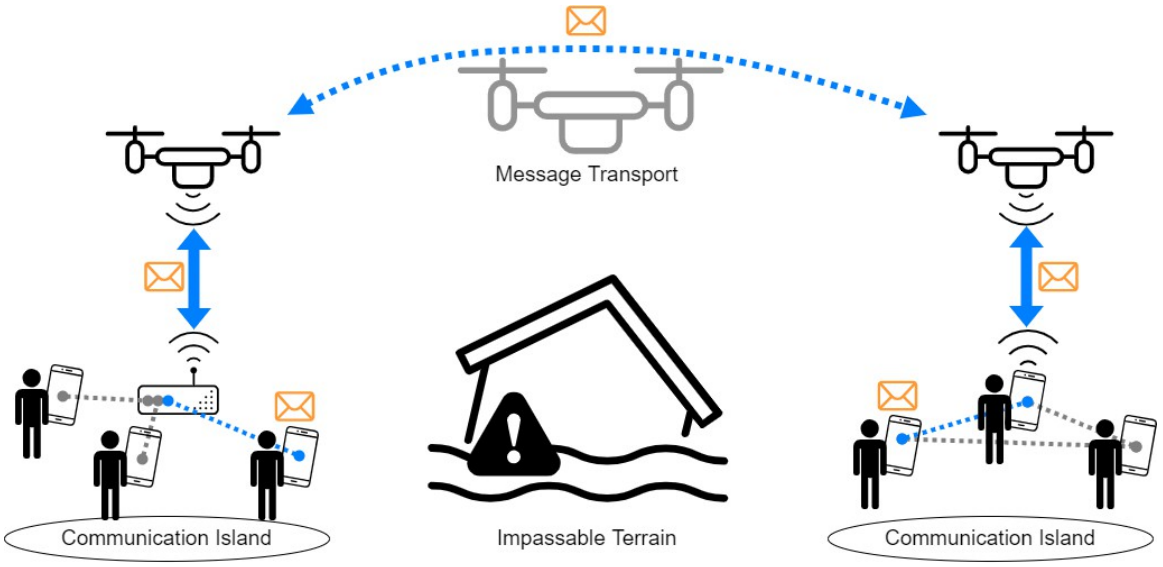


Figure 1. A communication network is separated by impassable terrain, e.g., debris or a flooded area, and traditional ICT infrastructure is not available, resulting in two separated communication islands. The utilization of UAVs as data ferries between the communication islands, independent to the blockades on the ground, provides delayed communication for affected civilians. However, this is only possible when the locations of the communication islands are known to the UAV system. (Flood Icon by Mapps from The Noun Project)

One of these proposed methods to quickly re-establish communication in a disaster area, is the utilization of small, cheap, off-the-shelf *Unmanned Aerial Vehicles* (UAV) as part of a larger aerial communication system. Communication capabilities for first responders and the civilian population can then be restored by providing, for example, a provisional Radio Access Network (RAN) (Chakraborty et al., 2018; Esrafilian et al., 2020; M. Gupta & Varma, 2021; Mahoro Ntwari et al., 2021; Moradi et al., 2018). Such an aerial RAN requires to place a significant number of UAVs over the disaster area to achieve sufficient coverage, demanding a vast amount of resources. Alternatively, UAVs can be used as long-distance data ferries between disconnected communication islands (Bujari et al., 2017; L. Gupta et al., 2016; Lieser et al., 2019; Mozaffari et al., 2017; Zeng & Zhang, 2017), as sketched in Figure 1. This approach requires physical message transport and, thus, provides only delayed message delivery, but is efficiently applicable with only a few available UAVs. In both cases, the provision of communication in the disaster area is paramount for disaster relief efforts and must be executed in the best possible way. Thus, a large part of research focuses on the very important aspects of efficient routing and optimal placement of UAVs. But both problems require precise and up-to-date information on the network topology to find a viable long-term solution. For example, topology information allows to identify communication gaps, which is needed for an initial establishment of routes or finding a suitable placement for access points. Furthermore, the topology in the disaster area may change over time, which requires updates on the topology to enable system adaptivity. The problem is that most research assumes the existence of such information and neglects its active collection, despite the fact that the availability and the acquisition of information from within a disaster area and its transfer to important stakeholders is not an easy task (Hagar, 2015).

In recent years, it was shown that UAVs can be used for the identification of networked communication devices on the ground, e.g., by WiFi or cellular signals (Rubina et al., 2019; Zeng et al., 2019). This ability is utilized in *Aerial Monitoring Systems*, which deploy UAVs to traverse and scan the disaster area to collect information on the network topology. The information is transferred to and collected at the base station, where it is combined and provided to different services, such as rescue services and disaster relief teams working in the area, or a UAV-based communication support system as described above (Arafat & Moh, 2021; Zobel et al., 2022). Similar to visual observation with cameras, signal-based identification and monitoring has a limited range. Thus, this traversal and scanning process of the UAVs is time-consuming, generally increasing in effort with larger disaster areas. Furthermore, the network topology can change over time. As a result, the quality of already gathered information will deteriorate over time without updating it. Therefore, the process must be repeated regularly, resulting in a constant monitoring mission for the *Aerial Monitoring System* (Zobel et al., 2022).

One of the most important aspects for the performance of an Aerial Monitoring System is the transfer of collected monitoring information from the UAVs to the base station. It is, however, also one of the less-researched challenges in this field. In the simplest solution, monitoring data is carried on the UAVs until their return to the base station. There, it is transmitted using physical connections or short-range high-throughput options like WiFi or Bluetooth (Zobel et al., 2022). Carrying collected information back to the base station may even be the only viable option, e.g., when cellular networks are destroyed in the disaster area. Additionally, most cheap, off-the-shelf UAVs do not have cellular or satellite communication interfaces. Due to the possibly long traversal times of a UAV while monitoring an area, carrying information back to the base station comes with the immense drawback of a high Age-of-Information on arrival. This may result in inaccurate or already wrong information given to the base station. Naturally, directly sending monitored information to the base station is a much better choice. But this requires the UAV to have a suitable communication channel, being adequate both in *data rate* to transmit the information fast and in *communication range* to reach the base station in the first place. Without cellular or satellite networks available, the question arises which alternatives for such a data channel can be used instead. New technologies like LoRa, providing robust communication with long communication ranges and low energy consumption, are promising candidates for infrastructure-independent communication in disaster areas. However, this comes with the drawback of small data rates, resulting in a low-throughput channel. Thus, UAVs may be able to reach their base station but probably not to send all monitored information — which requires a tradeoff solution that is yet not researched.

Therefore, this paper investigates the requirements of a suitable data channel for Aerial Monitoring Systems. We focus on the applicability of low-throughput channels like LoRa and possible methods to transmit monitored topology information from a disaster area. Specifically, we show that low-throughput channels are not suited to transfer typical topology information gathered by UAVs or an Aerial Monitoring System. Thus, we propose different concepts for structuring and aggregating network topology information, suitable for monitoring applications and for transmission in low-throughput data channels, respectively. Our concepts and aggregation methods are implemented and simulated in a generalized post-disaster network environment based on previous work (Zobel et al., 2022). With an in-depth evaluation of our concepts, we discuss possible combinations of data aggregation methods and communication channel properties that are suitable for direct data transmission in Aerial Monitoring Systems between monitoring UAVs and base stations on the example of long-range low-throughput communication technologies like LoRa. We show that by using our methods, aggregated monitoring information can be transferred to the base station with reasonable delay and reliability, even with a low-throughput channel bit rate as small as 300 $\frac{bit}{s}$. At the same time, the aggregated information still maintains an adequate representation of the monitored cluster topology of the disaster network. Our approach provides up-to-date information on the network topology, which is critical, e.g., for efficient disaster relief efforts and the resilient provision of a UAV-based communication service in the disaster area to the affected population.

The remainder of this paper is structured as follows. After the introduction and an overview of related topics given in this section, we discuss enablers for data transmission in low-throughput wireless networks as well as concepts for data aggregation and transmission handling. After that, the simulation scenario and the evaluation results are presented in detail, specifically focusing on channel requirements, message sizes, receiving delay, and the quality of representation of topology information from our proposed information aggregation methods. Finally, the paper is summarized and concluded.

ENABLING MONITORING INFORMATION TRANSMISSION FOR LOW-THROUGHPUT NETWORKS

The effective transmission of monitored topology information from Unmanned Aerial Vehicles (UAVs) to their base station is a critical aspect of every Aerial Monitoring System in disaster areas. In this paper, we focus on a post-disaster scenario without working ICT infrastructure, in which a fragmented disaster network based on short-range decentralized device-to-device communication is used (Á lvarez et al., 2018; Baumga¨rtner et al., 2020). The task of the deployed Aerial Monitoring System is to initially detect and subsequently monitor the topology of this disaster network. This information is highly relevant, e.g., for emergency services or UAV-based communication systems, to understand the location and distribution of civilians within the disaster area and efficiently provide support. The areas are scanned using a simple Boustrophedon-style route planning, as sketched in Figure 2. This back-and-forth algorithm is a well-known approach to address coverage path planning problems with deterministic area coverage (Choset & Pignon, 1998; Zobel et al., 2022). Within the scope of this paper, we do not research route planning or how information on a ground-based disaster network is collected in detail. The interested reader is referred to prior work (Zobel et al., 2022) that delves into these topics. Instead, we specifically focus on the transmission of information to the base station *after* its collection on a monitoring UAV. The basic information collected on a UAV consists of the location of a monitored device, a corresponding timestamp of the reading, as well as the network ID of the device itself.

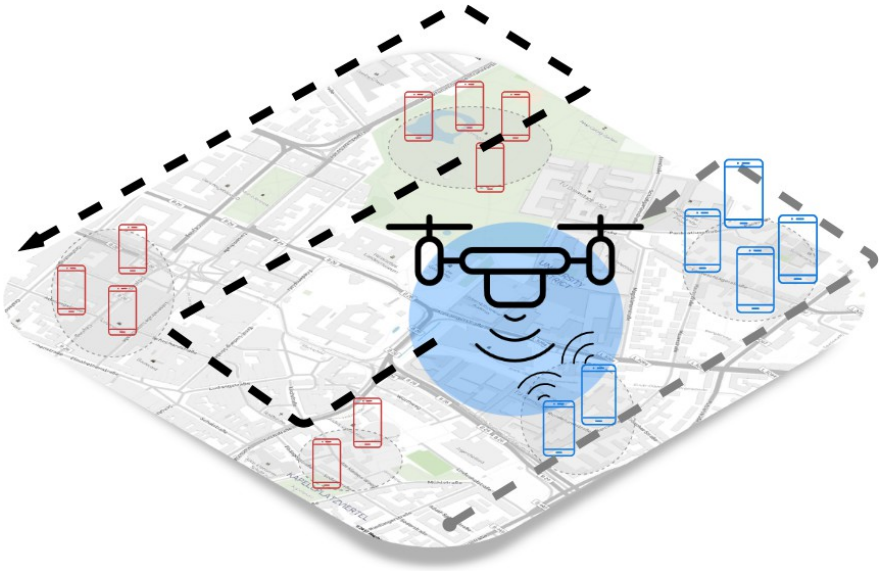


Figure 2. UAV-based detection of mobile devices in a disaster area. Using a Boustrophedon-style coverage path, the entire area is scanned. Devices within reception range of the UAV are detected (blue) by their signals. Due to extensive traversal routes of the UAV, the presence of undetected devices (red) may be unknown for a long time. Similarly, already seen devices can change their position without knowledge of the UAV.

One of the key characteristics of the considered disaster scenario is the highly fragmented network topology of distributed disaster networks (Á lvarez et al., 2018). Each of these fragments consists of individual devices that are connected with each other directly or over a multi-hop connection, but not with nodes of other fragments. Thus, we also speak of individual network clusters or communication islands. Due to the distributed nature of the disaster network, however, a monitoring UAV typically encounters individual nodes, not a cluster as a singular entity. Specifically, individual beacon messages of nodes are received while the UAV flies over the cluster. The UAV has to reconstruct the cluster topology based on connectivity estimations or additional information in beacons. For the former, a distance variable D_{conn}, describing the maximum connection distance between ground nodes, can be used to estimate cluster adherence. For the latter, beacon messages can include further information by which ground nodes themselves identify their affiliation to a cluster, e.g., a hash of all known messages (Khelil et al., 2007) or all known cluster nodes (Zobel et al., 2022). Naturally, data that is directly provided by the disaster network itself gives more accurate information, and thus, also a better representation of the actual network topology than a connectivity estimation.

Another limitation in the disaster scenario is that conventional cellular or satellite networks cannot be used at all by UAVs. We expect both communication interfaces not to be available on small and cheap, off-the-shelf UAVs and especially cellular networks to be impaired or entirely not available in the disaster area. However, the LoRa technology[1], a wireless communication technology originally developed for Internet-of-Things (IoT) applications and Low-Power Wide Area Networks (LPWAN), has shown its applicability as an alternative communication technology in disaster networks in recent years (Baumga¨rtner et al., 2017, 2020). Due to its cost-effectiveness, energy efficiency, adaptivity, and simplicity, LoRa is also increasingly integrated into UAVs and Unmanned Aerial Systems (Ghazali et al., 2021; Saraereh et al., 2020). Nevertheless, the inherent limitation of LoRa lies in its low throughput that comes as a tradeoff with its robust signaling. For typical devices[2], the usable bit rates range depending on the chosen LoRa parameters between 92 bps[3] and 21.9 kbps[4]. Thus, when considering the worst case, it would take a UAV about 0.7 seconds to transmit only a single 8-Byte-location measurement to the base station. Furthermore, this example does not include overhead or other important information, such as the time of the measurement or network information from the encountered network devices. Another heavy restriction is the typical 1% duty cycle of the used frequency bands (European Telecommunications Standards Institute (ETSI), 2018).

[1]Although we use LoRa as a well-researched example in the scope of this paper, our concepts and results are not exclusive to the LoRa technology and could be beneficial in any low-throughput network.
[2]For example, an ESP32 device using the LoRa SX1262 module.
[3]Parameters: 62.5 kHz BW, SF 12, CR 4:8
[4]Parameters: 500 kHz BW, SF 7, CR 4:5

Therefore, the following sections investigate how the transmission process and the structure of topology information can be optimized, to enable its transmission in such low-throughput networks. The objective is to reduce the packet size while preserving the integrity and representativeness of the transmitted data, thereby optimizing the UAV communication process.

Delaying Transmission

Whenever topology information is collected by a monitoring UAV, the question arises at which time this topology information can or should be transmitted to the base station. Intuitively, received information should directly be forwarded to minimize its Age-of-Information (AoI) on reception. This comes especially in mind, in case that information about the entire cluster is shared directly by the disaster network, such as all nodes and their locations within (Zobel et al., 2022). But whenever the cluster changes during the overflight, sending another message would be necessary. The optimal time to send the message is when the overflight of the cluster ends. At that point, the UAV cannot receive any more updates from the cluster, the information is most accurate, and no redundant messages are sent. However, pinpointing this exact time may be impossible. For individual node detection, furthermore, a UAV cannot know when all nodes in a cluster are detected — if that is even possible due to its flight path. Starting with a single node, each measurement must be grouped correctly together with other nodes, that are estimated to be in the same cluster. If a node changes position and the cluster splits as a result, this must also be considered in the process. Directly sending every measurement or every update of a possibly detected cluster or of singular nodes may impose a high load on the communication channel. Waiting for changes for a certain time before sending only the most up-to-date information could, therefore, be highly beneficial.

To address this, a delay λ is introduced, measured for every cluster. Whenever information for a cluster is received at the UAV, this delay is reset. Only when no change in information for that cluster has happened for that time, the cluster information is sent to the base station. In general, a small value for λ reduces the AoI at the base station, but may increase channel utilization due to redundant update messages. A larger delay, on the other hand, reduces information quality due to a larger AoI, but may optimize channel utilization.

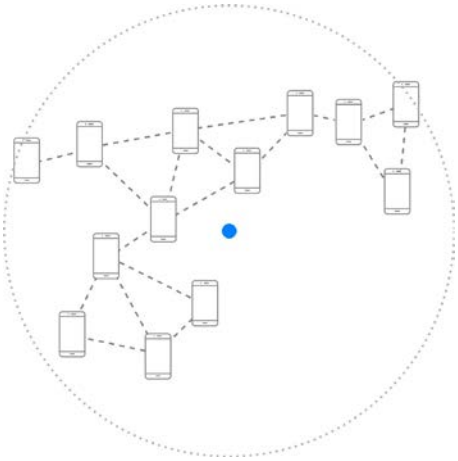
The specific value of λ could, for example, depend on the beacon send interval of the ground-based network. It determines at which rate information can be expected, usually in the range of a few seconds (Á lvarez et al., 2018). Whenever nothing is received after a multiple of that interval, information can be sent to the base. When this rate is unknown or cannot be determined by the UAVs, a static λ can be employed. Since typical smaller UAVs fly between 10 m/s to 30 m/s, λ must be chosen accordingly. As an example, $\lambda = 1s$ could be too low to overfly gaps in an oddly-shaped cluster and would lead to several necessary message transmissions, while $\lambda = 10s$ would result in possibly long-left clusters before sending the information. In general, λ must be chosen wisely due to the additional delay that worsens information quality.

The size of λ and its implications are addressed later-on. At this point, however, we note that the impact of λ on the age-of-information of monitoring data is tremendously smaller compared to a solution without a wireless data channel. A physical return for delivery can take a long time, as a single traversal of a monitored disaster area may take 30 minutes or more. Thus, any sensible choice of λ should perform better than a physical data delivery.

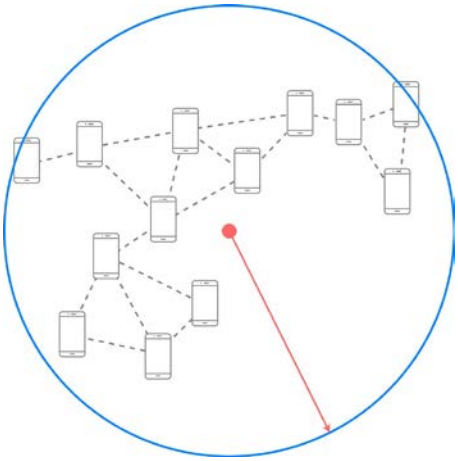
Aggregating Topology Information

To enable the utilization of low-throughput channels for topology information transmission, not only the number of messages must be reduced, but also their size. Within the scope of this paper, we are interested in monitoring the clustered network topology in the disaster area. Thus, the exact location of every individual node is not as important as the larger cluster topology they are part of. Furthermore, historic information is not as important as the most current topology. Therefore, only the most recent location of each node is interesting for the cluster topology.

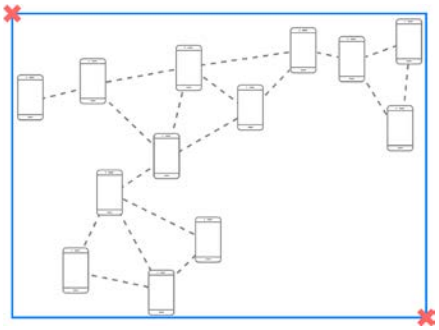
Instead of transmitting every monitored node location as a set of locations representing a cluster, we propose to aggregate the collected monitoring information directly on the UAV before sending. The goal is to reduce the size of the transmitted content as much as possible, while maintaining a suitable cluster representation. To address this, we specifically propose four simple geometric representations with different levels of aggregation: Point, Circle, Bounding Box, and Convex Hull. Figure 3 sketches out each representation with an exemplary disaster network cluster of mobile devices. Each approach provides a different accuracy in terms of representation of the actual cluster and different messages sizes for transmission to the base station. The aggregates are constructed based on the most recent information available to the UAV directly before transmission, i.e., only after the delay λ has expired, as described in the previous section.



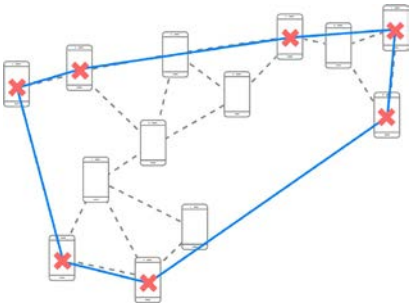
(a) **Point** — The center of the smallest enclosing circle is used as a single point to represent the cluster.



(b) **Circle** — The smallest enclosing circle, indicated by a location and a radius, is used to represent a cluster.



(c) **Bounding Box** — The bounding box, i.e., the smallest enclosing rectangle constructed by only two opposing corner vertices, is used as cluster representation.



(d) **Convex Hull** — The convex hull, i.e., the smallest convex shape enclosing the cluster, is used as representation. The number of required vertices depends on the cluster's shape.

Figure 3. Different cluster representation methods to aggregate topology information and to decrease its transmission size. Blue denotes the representing geometrical shape, red the required information that is transmitted for its representation. The exemplary cluster consists of mobile devices forming an ad hoc disaster network.

Point The Point representation aggregates all cluster information, visualized in Figure 3a by a disaster network cluster of mobile devices, into a single location. This point is calculated as the center of a smallest enclosing circle of all known cluster nodes. Point has the most reduced transmission size with a single location, e.g., being 8 Bytes in size. But this approach also provides the most inaccurate representation.

Circle The Circle representation (Figure 3b) extends Point by the radius of the smallest enclosing circle of the cluster. While increasing the transmission size, assuming 8 Bytes for the center location and another 4 Bytes for the circle radius, this provides a better representation of the cluster and its extents than a single point.

Bounding Box The Bounding Box representation (Figure 3c) uses two vertices to define the diagonally opposed corners of the smallest rectangle, in which all cluster nodes are contained. We align the bounding box in cardinal direction, thus, the horizontal edges are aligned West/East while the vertical edges are aligned North/South, respectively. This representation requires two locations, e.g., each 8 Bytes in size.

Convex Hull The Convex Hull approach, shown in Figure 3d, provides the most accurate cluster representation. Nevertheless, the transmission size can vary greatly, depending on the shape of the convex hull defined by its vertices, each being an 8-Byte location. More complex shapes require more vertices to form the convex hull, thus, also increasing the transmission size significantly.

Increasing Descriptiveness

The proposed approaches can tremendously reduce the transmission size compared to sending all information as individual measurements, especially for large clusters. Nevertheless, the descriptiveness of such detailed information is also lost by reducing it to geometric shapes: information on how many or which nodes are in a cluster, or how old a measured datum is not included in the aggregation from the beginning. We propose to include a single timestamp as minimal information for each aggregate, representing the oldest measurement in the aggregated data set. With this, the receiving base station is, at least, provided with a rough estimate of the data quality and can also differentiate information received at different times or from different sources more easily.

Furthermore, in case that a sufficient data channel is available, the aggregated cluster representations could again be enriched with additional information in exchange for a larger transmission size. As a first option, we can include the number of nodes represented by an aggregate. This only adds a simple and small integer value, but provides a clearer notion of the cluster size, which is valuable information, e.g., for disaster relief efforts to assess the importance of a cluster. However, a requesting service of our Aerial Monitoring System could require individual node information. In such cases, the network IDs of the enclosed nodes in the aggregated representations can be included as a second alternative. Although a valid option, this significantly increases the transmission size again, contrasting our initial goal of reducing it to enable usage of low-throughput networks. Nevertheless, this approach could be a reasonable tradeoff whenever network conditions are just slightly below the requirements for transmitting the entire topology in detail, but adequate with minor reductions.

EVALUATION

The evaluation is conducted in the Simonstrator (Richerzhagen et al., 2015) simulation framework, operating in the extended simulation environment designed for Unmanned Aerial Systems and ground-based Delay-Tolerant Networks (DTNs), as detailed in Lieser et al., 2019. This paper puts special emphasis on the less-researched communication between monitoring UAVs and base station. Thus, we utilize an available implementation of an Aerial Monitoring System from prior work (Zobel et al., 2022), which already includes monitoring UAVs and data collection of a disaster network. The existing implementation is extended by a LoRa channel with configurable data rate. Furthermore, the proposed topology information aggregation approaches and message additions discussed in the previous section are integrated.

Table 1. Simulation Environmental Settings.

Scenario	Map	Inner City, Post-Disaster
	Size	2000 m x 2000 m;
	Node Movement	Civilian Disaster Mobility [a]
	Points of Interest	5, random distribution
	Duration	3 h, 10 random seeds each
Disaster Network	Size	100 Nodes
	PHY	WiFi, IEEE 802.11g
	Range	approx. 75 m
	Data Rate	5 Mbit/s
Aerial Monitoring System	UAV	Quadrotor [b]
	Coverage Path Planning	Lawnmower
	Network Interaction	Cooperative Data Collection [c]
	Monitoring Mission Time	approx. 36 minutes
Monitoring Data Transmission	Aggregation	[None (Cluster), Point, Circle, Bounding Box (BBox), Convex Hull (Hull)]
	Additional Information	[∅, #Nodes, IDs]
	Transmission Delay λ	[None (λ_0), 5s, 15s, 30s, 60s]
	Channel Data Rate	[300 bps, 1 kbps, 21 kbps, 1 Mbps]

[a] cf. Zobel et al., 2021, [b] cf. Zobel et al., 2019, [c] cf. Zobel et al., 2022

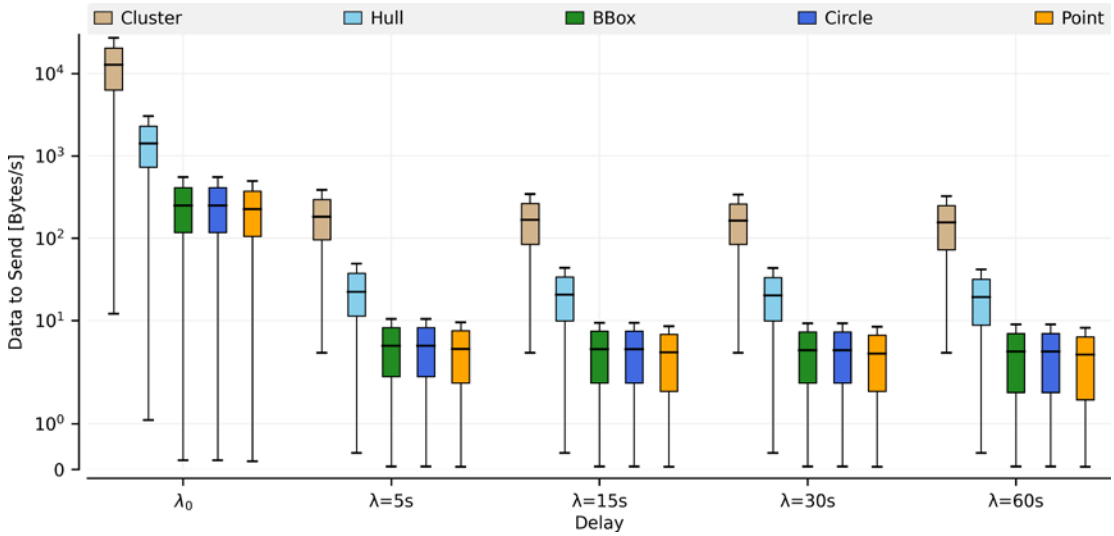
The simulation environment settings, listed in detail in Table 1, encompass 100 mobile devices moving in groups within an inner-city disaster area of 2x2 km^2 (Á lvarez et al., 2018; Zobel et al., 2021). For each simulation run of three hours, five points of interest like shelters and first aid stations are randomly selected, with the constraint that they must not overlap and must maintain a minimum distance of at least 300 m from each other. Devices predominantly gather around these points of interest and move in-between them. The mobility model relies on available Open Street Map (OSM) data, limiting device movement to streets and pedestrian walkways mapped in the OSM dataset. Each combination of simulation parameters is repeated ten times with different random seeds.

The entire area is monitored by a single UAV, completely traversing the monitoring area before returning to its base station. Traversal route planning uses a Boustrophedon-style coverage path planning approach (Choset & Pignon, 1998; Zobel et al., 2022), resulting in a total flight time of around 36 minutes from takeoff to landing. Topology aggregation can be switched between the four proposed aggregation methods Point, Circle, Bounding Box, and Convex Hull. Furthermore, the entire cluster information can be send without any aggregation (Cluster). Optional information in aggregates can be omitted (∅), include the number of nodes (#Nodes), or contain a list of all nodes' network IDs (IDs). The transmission delay λ, as described in the previous section, is investigated for no delay (λ_0), reasonably small values of $\lambda = 5s$ and $\lambda = 15s$, and much larger values of $\lambda = 30s$ and $\lambda = 60s$, respectively. Most importantly, we adjust the data rate of the transmission channel towards the base station to investigate requirements on it depending on the used aggregation methods and included detail of information in topology messages. Specifically, we test three settings that represent a data rate similarly achievable by using LoRa: A data rate of 300 bps constitutes the lower bound of possible data rates, 1 kbps a median tradeoff between throughput and signal robustness, and 21 kbps constitutes an upper bound. An additional data rate of 1 Mbps was evaluated as reference comparison for a generic cellular network connection.

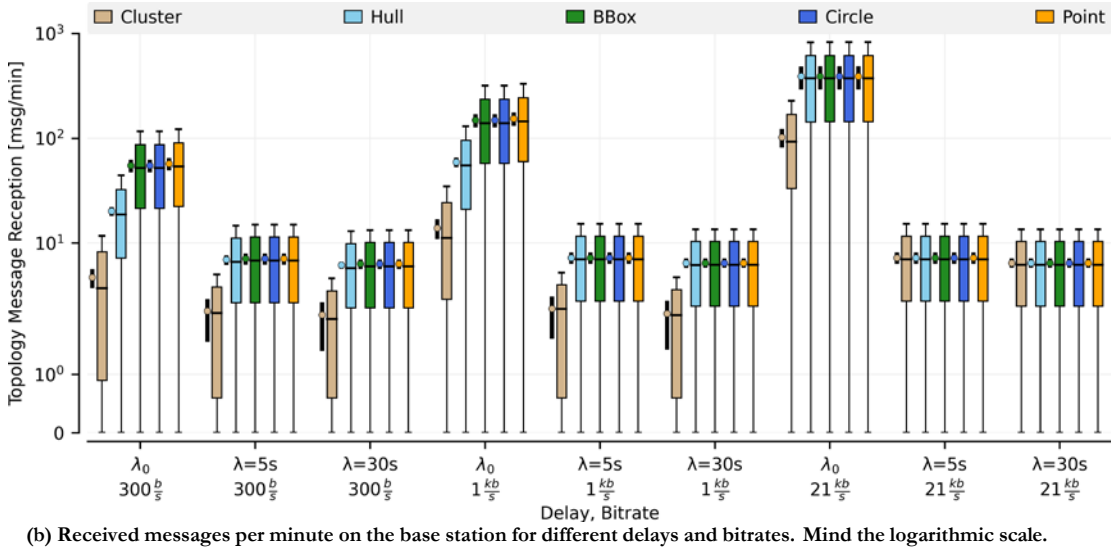
Sender Demands and Actual Throughput

At first, we take a look at the requirements of a monitoring UAV on the channel capacity for sending its information w.r.t. the transmission delay and the aggregation approach. When considering the entire monitoring mission time, the UAV queues between 2.9 and 13.8 messages per second when using direct transmission (λ_0) for around 75% of the time, which is independent of the aggregation approach. When using a transmission delay, however, this significantly reduces to a median of around 0.12 messages per second and to a maximum of 0.25 messages per second. With that, we have a first clue on the expected impact of using a transmission delay and not directly relaying all information received. Additionally, this clearly impacts the amount of data a monitoring UAV wants to send, as visualized in Figure 4a. For direct transmission, sending all information without aggregation would require a median of 13.4 $\frac{kByte}{s}$ but, e.g., only 194 $\frac{Byte}{s}$ with $\lambda = 5s$. At the same time, using aggregation with λ_0 already reduces the amount of data significantly, to around 1.5 $\frac{kByte}{s}$ using convex hulls and even further to around 250 $\frac{Byte}{s}$ for the other approaches. The differences between the transmission intervals, on the other hand, are only in the range of a few bytes per second, slightly decreasing with a higher interval. This suggests that only a few retransmissions were necessary, due to seen topology updates after sending the initial packet. The slight differences between the intervals also suggest that the shorter intervals are already sufficient to track a cluster's topology without requiring multiple topology messages to send during an overflight. Overall, the reduction in both the transmission delay and the aggregation approaches clearly highlights that information can be reduced on the sending side, which tremendously reduces the requirements for the data channel.

Next, we take a closer look at message reception. This clearly depends on the throughput of the channel: with a smaller channel, less messages can be sent and received per time unit. This bottleneck brings several problems, that are (i) an increase of the probability that messages are waiting longer in a sending queue, (ii) that messages are dropped from the queues due to expired waiting times, or (iii) that the queues are already full and new messages are dropped instantly. In the end, these problems are recognized on the receiver, as fewer messages are received than queued to send. Figure 4b visualizes the measured message reception in messages per minute on the base station. As expected from the reduction of sent messages seen in Figure 4a, using a transmission delay obviously decreases the number of messages on the receiving side. However, a larger transmission delay does not impact that number in a more significant way than a smaller delay. In contrary, — especially for the slower bitrate settings on the channel — the number of messages for Cluster and Convex Hull are *lower*, despite a higher number of messages queued on the sender. This already suggests that messages are dropped. With an increase in bitrate, on the other hand, the number of received messages increases for λ_0 for all approaches, while leveling up for Cluster with the other aggregated approaches for a bitrate of 21 $\frac{kBit}{s}$ using a transmission delay. Thus, the number of message drops reduces generally with a higher bitrate, as expected. More importantly, a bitrate of 1 $\frac{kBit}{s}$, which is achievable by state-of-the-art LoRa modules, shows to be sufficient to transmit the aggregated topology information without perceivable message drops.



(a) Throughput demands for sending packets on an unlimited monitoring UAV with different delays. Mind the logarithmic scale. While we perceive a drastic decrease in demand between sending without and with delay, the influence of different delays is minimal.



(b) Received messages per minute on the base station for different delays and bitrates. Mind the logarithmic scale.

Figure 4. Throughput demands for an unlimited sender compared to the received messages on the base station with limited bitrates.

The loss of messages between sender and receiver is further investigated considering Figure 5. As before, a clear decrease in dropped messages with increasing bitrate and with using topology aggregation is perceived. In comparison to direct transmission, only a fraction of messages are dropped due to the smaller number and size of messages, even with a bitrate of only 300 $\frac{Bit}{s}$. These results also support the findings that a data rate of 1 $\frac{kBit}{s}$ is already sufficient for the transmission of aggregated topology information without perceivable losses due to channel capacity. For the Cluster approach, however, this still constitutes for a message loss of around 70%. In addition to the other figures, Figure 5 exemplary visualizes the results for 1 $\frac{MBit}{s}$ data rate to highlight that this is sufficient to send all information without aggregation and loss over the data channel.

Message Size and Optional Message Content

As depicted in Figure 6a, the size of messages is an important factor for their transmission in low-throughput channels. Especially for very large Cluster messages, but also seen in detail for larger Convex Hull messages, these are not received on the base station due to the channel limitations. First with a data rate of 21 $\frac{kBit}{s}$, and supported by the measurements at 1 $\frac{MBit}{s}$ when all messages are received without problems, very bulky messages can be transmitted. The large number of dropped Cluster messages highlights that they are too large in size to be

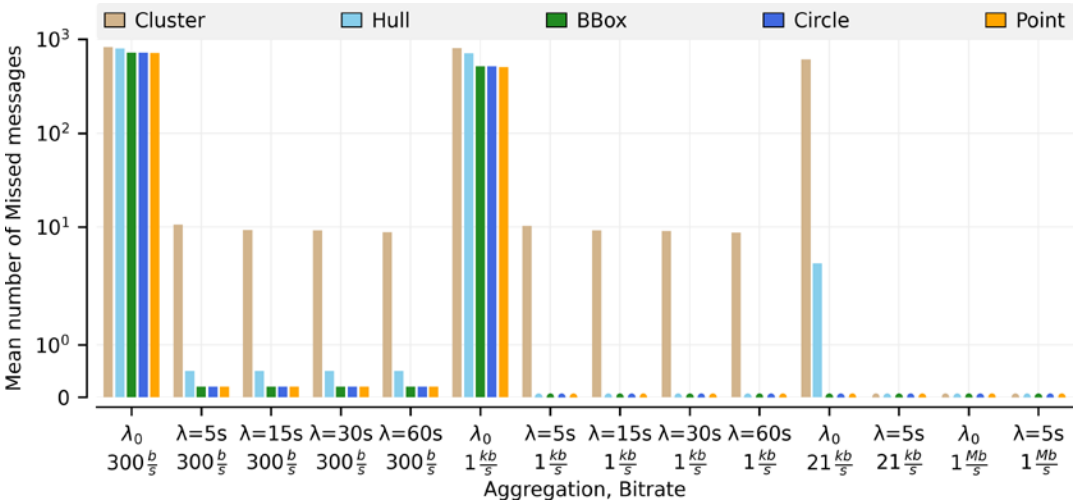


Figure 5. Mean number of missed messages, aggregated over all simulation runs. Mind the logarithmic scale. Round caps indicate zero values. With increasing bitrate and by using topology aggregation, the rate of dropped messages decreases significantly. For aggregated topology information, a data rate of 1 $\frac{kBit}{s}$ is already sufficient for transmission.

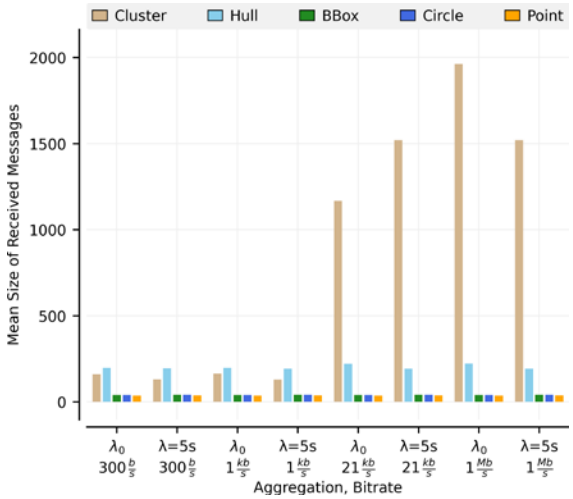
transmitted successfully in low-throughput networks. Thus, the proposed aggregation of information and reduction in size is clearly necessary to allow data transmission for larger clusters in the first place.

In the previous section, we also introduced optional message content in the information aggregates, to increase their expressiveness towards the receiving base station. But we also discussed that, especially including a list of all node IDs, significantly increases the message size, contradicting the intended size reduction. Therefore, Figure 6b takes a closer look of the influence of the optional message content in aggregate messages, on the example of 1 $\frac{kBit}{s}$ bitrate. Clearly, there is no impact by adding a simple integer value for the number of represented nodes in the aggregates, which is also true for a bitrate of 300 $\frac{Bit}{s}$ and the higher bitrates, respectively. For the list of IDs, however, an increase of messages drops is perceivable due to the comparably extreme message size of aggregates. They can increase to more than ten times the size of the same message without these IDs. This increase is especially problematic for the lower bitrate channels, since message drops re-appear for aggregated topology messages with a formerly sufficient bitrate of 1 $\frac{kBit}{s}$ and strongly increase for a bitrate of 300 $\frac{Bit}{s}$. This supports our statement that the list of message IDs should only be added when absolutely necessary and with a sufficiently scaled channel throughput. Preventing message loss and sending the aggregated topology information to the base station is more important than this optional information.

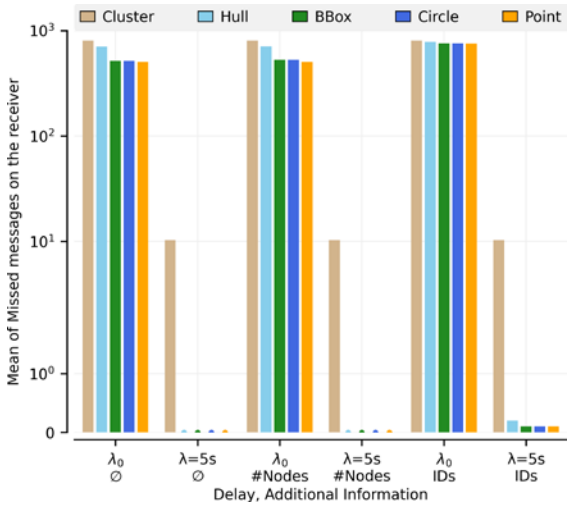
Transmission Time

Another problem with limited channel throughput is the additional delay introduced by filled sending queues. In addition to the actual airtime of a packages, i.e., the time to send the entire package from start to end on the wireless medium, this results in a certain transmission time per packet. To assess this transmission time, we traced the time difference between putting a packet in the network interface's queue on the monitoring UAV and receiving it at the network interface on the base station. This transmission time is visualized in Figure 6c.

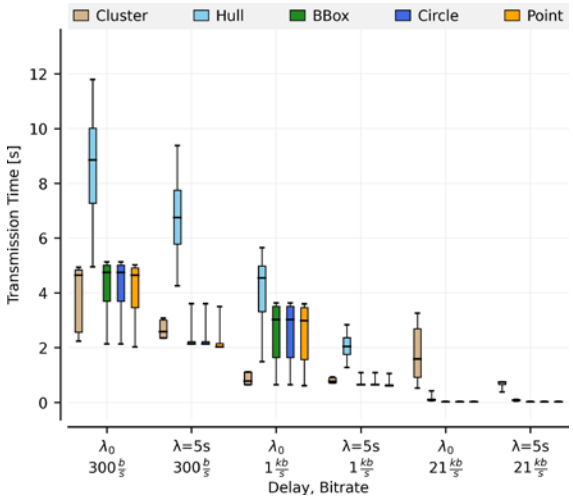
Directly perceivable is that transmission times are generally higher for direct transmission than for a delayed transmission. This comes as a direct result of longer sending queues, which increases the delay for all messages. Furthermore, messages for Convex Hull are clearly more delayed than other types. This can be attributed to their significantly larger size compared to the other aggregation approaches. Therefore, they take longer to transmit and extend the waiting time in the queue for all messages. For the other aggregation approaches, only slight differences can be seen from the results. Curiously, transmission times for Cluster messages are much lower for 300 $\frac{Bit}{s}$ and 1 $\frac{kBit}{s}$ bitrate. Nevertheless, Figure 6c only depicts the time for *received* messages — and, as seen before, these are only the small ones. Larger messages are only received for 21 $\frac{kBit}{s}$. With that bitrate, the transmission time increases again because most messages are now countable for the transmission time. Similarly, nearly no delay is introduced for the aggregated approaches by the queues, indicating empty queues and nearly instantaneous transmission.



(a) Average size of received messages per bitrate and delay transmission. With a higher bitrate, very large Cluster and Convex Hull messages can be transmitted.

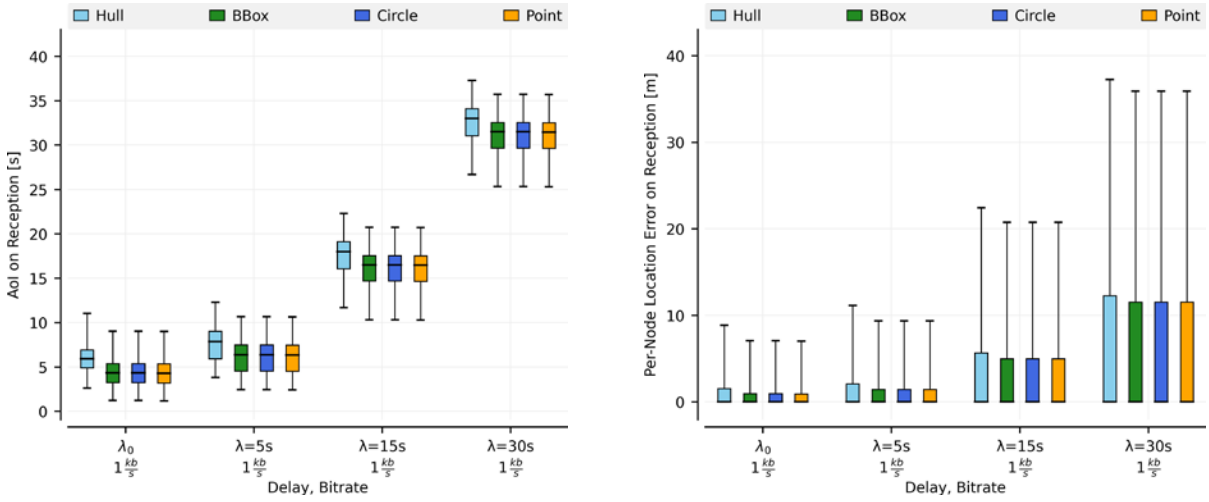


(b) Missed Messages per optional message content, using a data rate of 1 kbps. Using aggregation, only Convex Hull results in a few message drops for ∅ and #Nodes. When sending a list of all node IDs, however, the number of message drops increases intensely for direct transmission, and slightly for $\lambda = 5s$.



(c) Transmission time as difference between message insertion at the sending queue on the UAV and message reception at the base station. Only received messages can be counted with this metric, which must be considered when assessing this time difference especially for larger messages and smaller bitrates.

Figure 6. Received message size, missed messages on the receiver for optional message content, and transmission time. The results highlight that the proposed aggregation methods enable transmission via low-throughput data channels in contrast to sending all data.



(a) The age of monitored information when received on the base station, measured per node information aggregated on the monitoring UAV.

(b) The location error of information received on the base station, measured per node information aggregated on the monitoring UAV.

Figure 7. Per-node Age-of-Information and absolute location error on reception for a data rate of 1 $\frac{kbit}{s}$.

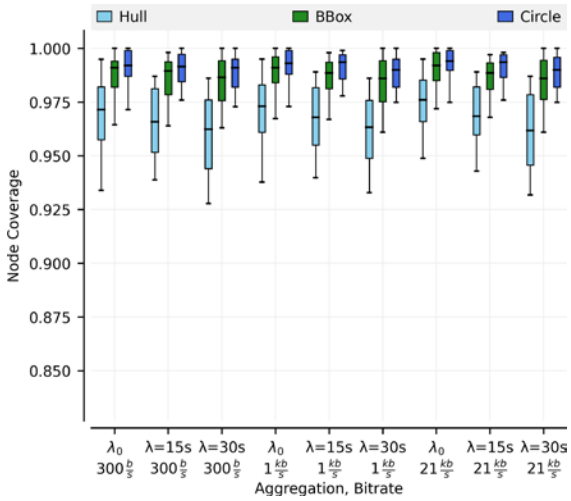
Topology Representation of Aggregations

The remaining question is, whether the aggregated information in the form of the proposed simple geometric shapes is suitable to represent the actual topology that was monitored by the UAV.

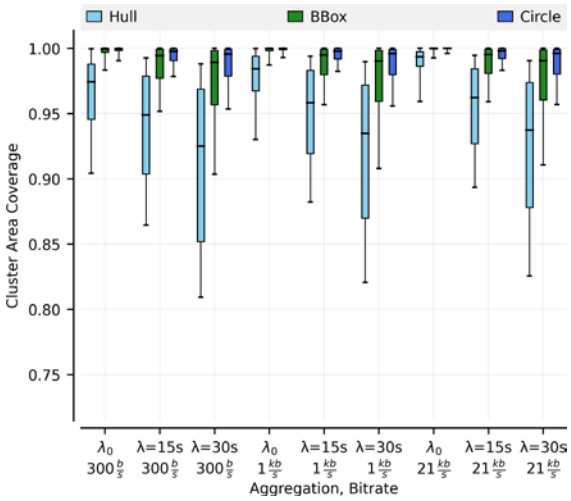
First, we look at the Age-of-Information (AoI) and the location error (LE) of the information. Both metrics encompass the difference from the initial measurement on the UAV until the final reception and assessment on the base station. Therefore, they provide a possibly expected error margin, introduced by a delayed transmission ($\lambda > 0$) and waiting times due to message queues and limited channel throughput. Figure 7 depicts results for both metrics on the example of a 1 $\frac{kbit}{s}$ bitrate. As expected, we see a direct correlation between the transmission delay λ and the AoI. However, we also see a generally larger AoI for the Convex Hull approach due to longer transmission time required as a result of the larger message sizes. Since the AoI can directly influence the location error, as it gives more time for the monitored devices on the ground to possibly relocate, the LE similarly increases with a higher delay. Nevertheless, the median LE is still zero, providing an acceptable information quality, at least within the simulated scenario. However, these results may change with other scenarios or simulated mobility models. In general, a low AoI and location error is preferable; the transmission delay should be kept as small as possible.

Second, we specifically regard the representation and accuracy of the aggregated topology information compared to the actual network topology. Figure 8 provides an overview of the used metrics. As geometric comparison baseline serves the convex hull of the network cluster at the time of receiving the aggregate at the base station. Since the Point approach does not provide an actual geometric area, no comparisons are possible.

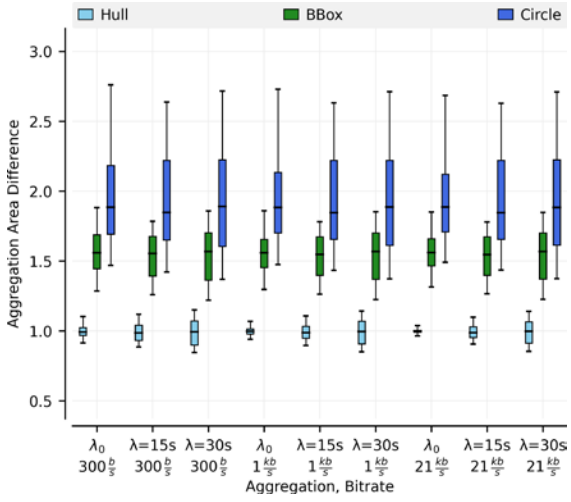
Most importantly for our approaches is, that they cover the network clusters constructed from individual nodes to represent the actual network topology. As depicted in Figure 8a, the node coverage of Convex Hull, Bounding Box, and Circle aggregates is very good, with more than 92.5% of nodes always covered. Furthermore, the less exact approaches Bounding Box and Circle usually cover more nodes than Convex Hull. This results from their larger but less precise extents, allowing more node movement on the ground before losing coverage. On the other hand, there are no significant differences on node coverage for different transmission delays. Additionally, there are only slight performance increases with a higher data rate. For the coverage of the geometric cluster representation, however, the influences of mobility in the network are much clearer. Figure 8b highlights a decrease in cluster coverage with a higher transmission delay and better coverage with higher data rate. Similar to node coverage, the less-exact Bounding Box and Circle generally cover more of the clusters, even with mobility, than the more exact Convex Hull. The tradeoff, however, becomes clear when comparing the difference in size between the aggregation and the actual network cluster (cf. Figure 8c) as well as the overlap of the aggregation with the network cluster (cf. Figure 8d), respectively. The circular aggregation area is between 1.25 and 2.75 times bigger than the actual cluster, including only 40% to 70% of it. The bounding box representation performs better, but still depicts an area up to 1.8 times the size of the actual cluster. Only the Convex Hull approach is able to closely represent the network clusters. But with a higher susceptibility to node mobility, the transmission delay also has a much larger impact on its accuracy than on the other approaches.



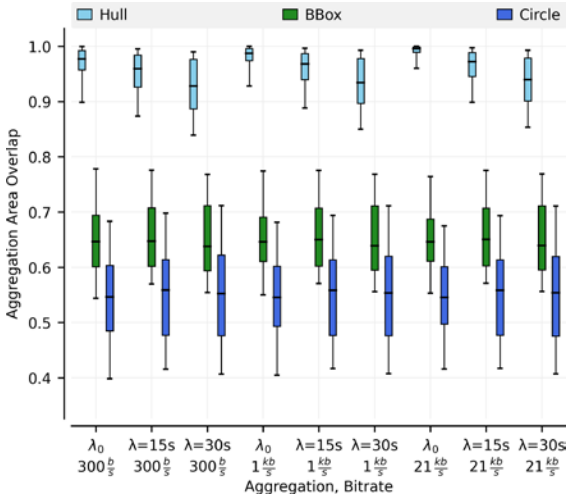
(a) Node coverage.



(b) Cluster area coverage.



(c) Aggregation vs. actual cluster size.



(d) Part of aggregation overlapping with actual cluster.

Figure 8. Overview of comparing metrics for the representation of aggregated topology information and the actual network topology and network clusters on the ground. Note that the y-axes are not starting at zero. For comparisons with the actual cluster area, the convex hull of the network cluster is used.

CONCLUSION

The possible impact and the necessary requirements of a data channel between Unmanned Aerial Vehicles (UAVs) and their operating base station is one of the less-researched topics in the field of Unmanned Aerial Systems. Especially for low-cost off-the-shelf UAVs without proper network interfaces, such as for satellite or cellular networks, alternatives are required to allow their application in large-scale disaster areas, where a central ICT infrastructure may not be available.

This paper discusses the importance of such a data channel on the example of an Aerial Monitoring System, that searches and monitors mobile devices in a disaster scenario with UAVs. This information is highly valuable, for example, to rescue services, but may arrive only with severe delay without a proper data channel available to the monitoring UAVs. LoRa is one of the suitable wireless communication technologies that could be integrated and used on off-the-shelf UAVs and we are currently working on a possible hardware implementation. LoRa provides a very robust, but also very limited low-throughput communication channel. We have shown that the bitrates achieved by LoRa are not suitable to transfer typical topology information monitored by UAVs. Therefore, we propose to aggregate topology information as simple geometric shapes, which can be transmitted using only small, basic information like a few locations or numerical values. Our evaluation results demonstrate that theses aggregations significantly reduce the size of messages for transmission to the base station. At the same time, they maintain an adequate representation of the monitored clusters and the network topology. The aggregated monitoring information can be transferred with reasonable delay and reliability towards the base station even over low-throughput networks, tested in this paper with a bitrate as small as 300 $\frac{bit}{s}$. A higher throughput reduces transmission times and message drops, further increasing the usability of both our approach and the gathered monitoring information.

Future work must encompass more realistic representations of the data channels. In this work, we focused on the application of information aggregation and its usability within limited data channels, but neglected signal losses, packet collisions, and more. Furthermore, an increase of devices in the disaster network and multiple monitoring UAVs working in parallel may further burden the used data channel, reducing its usability and effective throughput. More sophisticated, possibly situation-adaptive aggregation approaches are, therefore, necessary in combination with larger and more complex network environments. Similarly, the reduction of transmission data could be approached as an optimization problem, where only the optimal subset of actual location measurements is sent to reduce channel utilization. Nevertheless, the work in this paper provides a foundation for future research in an important, but currently unattended and not well-researched subject of Unmanned Aerial Systems.

ACKNOWLEDGMENT

This work has been funded by the LOEWE initiative (Hessen, Germany) within the *emergenCITY* center and the German Research Foundation (DFG) within the Collaborative Research Center (CRC) 1053 MAKI.

REFERENCES

Á lvarez, F., Almon, L., Lieser, P., Meuser, T., Dylla, Y., Richerzhagen, B., Hollick, M., & Steinmetz, R. (2018). Conducting a Large-Scale Field Test of a Smartphone-Based Communication Network for Emergency Response. *Proceedings of the 13th Workshop on Challenged Networks*, 3–10.

Arafat, M. Y., & Moh, S. (2021). Bio-Inspired Approaches for Energy-Efficient Localization and Clustering in UAV Networks for Monitoring Wildfires in Remote Areas. *IEEE Access*, *9*, 18649–18669.

Baumga¨rtner, L., Kohlbrecher, S., Euler, J., Ritter, T., Stute, M., Meurisch, C., Mu¨ hlha¨user, M., Hollick, M., von Stryk, O., & Freisleben, B. (2017). Emergency Communication in Challenged Environments via Unmanned Ground and Aerial Vehicles. *Proceedings of the Global Humanitarian Technology Conference (GHTC)*.

Baumga¨rtner, L., Lieser, P., Zobel, J., Bloessl, B., Steinmetz, R., & Mezini, M. (2020). LoRaAgent: A DTN-Based Location-Aware Communication System using LoRa. *Proceedings of the 10th IEEE Global Humanitarian Technology Conference (GHTC)*.

Bujari, A., Calafate, C. T., Cano, J.-C., Manzoni, P., Palazzi, C. E., & Ronzani, D. (2017). Flying Ad-Hoc Network Application Scenarios and Mobility Models. *SAGE Publications Sage UK International Journal of Distributed Sensor Networks*, *13*(10).

Chakraborty, A., Chai, E., Sundaresan, K., Khojastepour, A., & Rangarajan, S. (2018). SkyRAN: A Self-Organizing LTE RAN in the Sky. *Proceedings of the 14th International Conference on Emerging Networking Experiments and Technologies*, 280–292.

Choset, H., & Pignon, P. (1998). Coverage path planning: The boustrophedon cellular decomposition. *Field and service robotics*, 203–209.

Esrafilian, O., Gangula, R., & Gesbert, D. (2020). Autonomous uav-aided mesh wireless networks. *Proceedings of the 2020 IEEE Conference on Computer Communications Workshops*, 634–640.

European Telecommunications Standards Institute (ETSI). (2018). ETSI EN 300 220-2: Short Range Devices (SRD) operating in the frequency range 25 MHz to 1000 MHz; Part 2: Harmonised Standard for access to radio spectrum for non specific radio equipment.

Fehling-Kaschek, M., Miller, N., Haab, G., Faist, K., Stolz, A., Ha¨ring, I., Neri, A., Celozzi, G., Sanchez, J., & Valera, J. (2020). Risk and Resilience Assessment and Improvement in the TeleCommunication Industry. *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference (PSAM15)*.

Gallucci, M. (2018). Rebuilding Puerto Rico's Power Grid: The Inside Story. *IEEE Spectrum*, *55*(5), 30–38.

German Federal Agency for Civic Education. (2021). Jahrhunderthochwasser 2021 in Deutschland (translation: Flood of the Century 2021 in Germany), (online) https://www.bpb.de/politik/hintergrund-aktuell/337277/jahr-hunderthochwasser-2021-in-deutschland, last accessed: 20.12.2023.

Ghazali, M. H. M., Teoh, K., & Rahiman, W. (2021). A systematic review of real-time deployments of uav-based lora communication network. *IEEE Access*, *9*, 124817–124830. https://doi.org/10.1109/ACCESS.2021.3110872

Gupta, L., Jain, R., & Vaszkun, G. (2016). Survey of Important Issues in UAV Communication Networks. *IEEE Communications Surveys & Tutorials*, *18*(2), 1123–1152.

Gupta, M., & Varma, S. (2021). Metaheuristic-Based Optimal 3D Positioning of UAVs Forming Aerial Mesh Network to Provide Emergency Communication Services. *Wiley Online Library IET Communications*, *15*(10), 1297–1314.

Hagar, C. (2015). Crisis informatics. In *Encyclopedia of Information Science and Technology, Third Edition* (pp. 1350–1358). IGI Global.

Hollick, M., Hofmeister, A., Engels, J. I., Freisleben, B., Hornung, G., Klein, A., Knodt, M., Lorenz, I., Lieser, P., Mühlhäuser, M., Pelz, P., et al. (2019). EmergenCity: A Paradigm Shift Towards Resilient Digital Cities. *Proceeding of the World Congress on Resilience, Reliability and Asset Management (WCRRAM)*, 383–406.

Khelil, A., Marro´n, P. J., Becker, C., & Rothermel, K. (2007). Hypergossiping: A Generalized Broadcast Strategy for Mobile Ad Hoc Networks. *Elsevier Ad Hoc Networks*, *5*(5), 531–546.

Kreienkamp, F., Philip, S. Y., Tradowsky, J. S., Kew, S. F., Lorenz, P., Arrighi, J., Belleflamme, A., Bettmann, T., Caluwaerts, S., Chan, S. C., et al. (2021). Rapid Attribution of Heavy Rainfall Events Leading to the Severe Flooding in Western Europe during July 2021. *World Weather Attribution*.

Lieser, P., Zobel, J., Richerzhagen, B., & Steinmetz, R. (2019). Simulation Platform for Unmanned Aerial Systems in Emergency Ad Hoc Networks. *Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM), awarded with the Best Student Paper Award*.

Mahoro Ntwari, D., Gutierrez-Reina, D., Toral Mar´ın, S. L., & Tawfik, H. (2021). Time Efficient Unmanned Aircraft Systems Deployment in Disaster Scenarios Using Clustering Methods and a Set Cover Approach. *MDPI Electronics*, *10*(4), 422.

Moradi, M., Sundaresan, K., Chai, E., Rangarajan, S., & Mao, M. Z. (2018). Skycore: Moving core to the edge for untethered and reliable uav-based lte networks. *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, 35–49.

Mori, A., Okada, H., Kobayashi, K., Katayama, M., & Mase, K. (2015). Construction of a Node-Combined Wireless Network for Large-Scale Disasters. *Proceedings of the 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*, 219–224.

Mozaffari, M., Saad, W., Bennis, M., & Debbah, M. (2017). Wireless Communication Using Unmanned Aerial Vehicles (UAVs): Optimal Transport Theory for Hover Time Optimization. *IEEE Transactions on Wireless Communications*, *16*(12), 8052–8066.

Mu¨hlha¨user, M. (2021). Panel: Resilience in Urban and Critical Infrastructures — The Role of Pervasive Computing. *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 1–2.

Ranghieri, F., & Ishiwatari, M. (2014). *Learning from MegaDisasters: Lessons from the Great East Japan Earthquake*. World Bank.

Richerzhagen, B., Stingl, D., Ru¨ckert, J., & Steinmetz, R. (2015). Simonstrator: Simulation and Prototyping Platform for Distributed Mobile Applications. *EAI International Conference on Simulation Tools and Techniques (SIMUTOOLS)*, 1–6.

Rubina, A., Artemenko, A., & Mitschele-Thiel, A. (2019). Towards a Realistic Path Planning Evaluation Model for Micro Aerial Vehicle-Assisted Localization of Wi-Fi Nodes. *Proceedings of the 44th IEEE Conference on Local Computer Networks (LCN)*, 169–176.

Saraereh, O. A., Alsaraira, A., Khan, I., & Uthansakul, P. (2020). Performance evaluation of uav-enabled lora networks for disaster management applications. *Sensors*, *20*(8).

Toya, H., & Skidmore, M. (2018). Cellular Telephones and Natural Disaster Vulnerability. *MDPI Sustainability*, *10*(9), 2970.

Universite´ Catholique de Louvain (UCL). (2022). EM-DAT: The Emergency Events Database.

Whittaker, J., Haynes, K., Handmer, J., & McLennan, J. (2013). Community Safety During the 2009 Australian 'Black Saturday' Bushfires: An Analysis of Household Preparedness and Response. *CSIRO International Journal of Wildland Fire*, *22*(6), 841–849.

Zeng, Y., & Zhang, R. (2017). Energy-Efficient UAV Communication with Trajectory Optimization. *IEEE Transactions on Wireless Communications*, *16*(6), 3747–3760.

Zeng, Y., Wu, Q., & Zhang, R. (2019). Accessing From the Sky: A Tutorial on UAV Communications for 5G and Beyond. *Proceedings of the IEEE*, *107*(12), 2327–2375.

Zobel, J., Kundel, R., & Steinmetz, R. (2022). CAMON: Aerial-Ground Cooperation System for Disaster Network Detection. *Proceedings of the 19th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*.

Zobel, J., Lieser, P., Drescher, B., Freisleben, B., & Steinmetz, R. (2019). Optimizing Inter-Cluster Flights of Post-Disaster Communication Support UAVs. *Proceedings of the 44th IEEE Conference on Local Computer Networks (LCN)*.

Zobel, J., Lieser, P., Meuser, T., Baumga¨rtner, L., Mezini, M., & Steinmetz, R. (2021). Modeling Civilian Mobility in Large-Scale Disasters. *Proceedings of the 18th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*.

Zorrilla, C. D. (2017). The View from Puerto Rico – Hurricane Maria and its Aftermath. *New England Journal of Medicine*, *377*(19), 1801–1803.

Location and allocation of primary and backup shelters in transboundary disasters

Oscar Rodríguez-Espíndola
Aston Business School, Aston University
o.rodriguez-espindola@aston.ac.uk

Pavel Albores
Aston Business School, Aston University
p.albores@aston.ac.uk

Dolores Luna-Reyes
Universidad de las Americas Puebla
dolorese.luna@udlap.mx

Juan Díaz-García
Universidad de las Americas Puebla
juana.diaz@udlap.mx

Sergio Picazo-Vela
Universidad de las Americas Puebla
sergio.picazo@udlap.mx

ABSTRACT

Disaster management is plagued by uncertainty and chaotic situations, often crossing boundaries and affecting multiple jurisdictions. Collaboration in these settings is complicated, especially when authorities face failure in the infrastructure available leading to facilities operating over capacity and/or vulnerable people left without support. There are records about those experiences globally. There is a broad literature on shelter location/allocation, but few articles consider the potential failure facilities and none of those contributions consider the collaboration between multiple jurisdictions. This article introduces a novel bi-objective formulation for shelter location-allocation considering backup facilities for operations involving multiple jurisdictions. This approach is applied to a set of numerical examples and a real case of volcanic eruptions in Mexico. The results show the influence of transboundary coordination on the support provided to disaster victims, the value of backup facilities and the capacity of the model design to provide relevant alternatives in practice.

Keywords

Shelter location-allocation; transboundary coordination; backup facilities; stochastic optimization.

INTRODUCTION

Just in 2022, disasters caused by natural hazards threatened over 185 million people and caused an estimated US $223.8 billion of economic losses (CRED, 2023). Humanitarian logistics has arisen as a critical field to develop solutions to handle the conditions created by disasters. It involves multiple stakeholders collaborating to deliver products and services to satisfy the needs of people affected by disasters (Gutjahr et al., 2020).

Evacuation and shelter selection influence the service experienced by people affected by disasters. Shelters are often prepared in advance, but emergency services and facilities can fail under disaster situations (Albareda-Sambola et al., 2015). Although there has been an increase of models incorporating backup facilities, there is limited information about the value of these facilities, especially when facility capacity is included in the formulations.

Humanitarian logistics is a multi-actor, multi-agency, and cross-jurisdictional activity involving private and public organizations (Oloruntoba, 2010). Despite the emphasis on the interaction between different stakeholders, coordination under disaster circumstances remains a key challenge (Balcik et al., 2010). This complexity is further

WiP Paper – Analytical Modeling and Simulation
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

fueled by the need to conduct joint operations for crises across boundaries affecting different jurisdictions (Boin et al., 2014) because each jurisdiction has defined responsibility and attributions.

Disasters can transcend political, functional and temporal boundaries, which is translated into significant complications to handle them requiring extreme adaptation (Ansell et al., 2010). Recent research highlights the importance of incorporating the collaboration of agencies from different jurisdictions (Nurmala et al., 2017; Olsson, 2015). However, that has not been reflected in the development of cross-jurisdictional models for disaster response. This study answers two research questions: i) what is the effect of coordination for shelter selection between different jurisdictions in the service provided to people affected? and ii) what is the value of incorporating backup shelters for disaster response?

This research proposes a two-stage stochastic optimization model to support evacuation and shelter location decisions incorporating cross-jurisdictional collaboration. It is tested using a set of numerical examples and a case study. It contributes to the literature by proposing a model for operational collaboration in shelter management between stakeholders from different jurisdictions, acknowledging the possibility of failure of any facility using backup facilities, and investigating the impact of transboundary coordination on the experience of the people affected by the disaster.

LITERATURE REVIEW

Backup Shelters

Shelter location involves the selection of suitable safe points to provide protection to people affected by disasters (Kınay et al., 2018). The vulnerability of facilities inside the affected area has motivated studies looking into the use of backup facilities for distribution centers (Mohamadi et al., 2019; Rahmani et al., 2018; Yahyaei & Bozorgi-Amiri, 2019). The need for backup shelters to react to failures in sheltering, however, has been understudied, with many studies neglecting to consider shelter failure (Ma et al., 2016).

There is a stream of articles focused on selecting the location for building shelters in suitable areas (Aghaie & Karimi, 2022; Geng et al., 2021), thereby reducing the possibility of shelter failure. However, in many countries the policy is to use existing facilities as shelters (Liu et al., 2011), which are vulnerable to failure. The work of Gangwal et al. (2023) and Wang and Xu (2023) shows that it is important to look at the vulnerability of the shelter not only because of damage on the facility itself, but also because of road failure leading to isolated shelters.

Albareda-Sambola et al. (2015) introduce the use of the p-next center problem to select primary and backup facilities closer to the victims in case they reach their closest center and it is closed. Albareda-Sambola et al. (2019) extend that idea by considering different strata of people requiring different services whereas Hinojosa et al. (2023) use a hierarchical process where the planner sets service centers based on the worst-case behavior of the public and the beneficiaries select the center based on their preferences. Although this is a promising avenue of research, all the models disregard the capacity of the facilities, which is a very practical constraint in reality. An et al. (2013) propose a formulation assuming each facility is subject to independent probabilistic disruptions which are represented in a set of scenarios. Their formulation defines the allocation upon the realization of the corresponding scenario, which requires the government to notify and direct all the evacuees. Aiming to maximize network reliability and considering facility capacity, Babashahi et al. (2020) use reserve hub elements as backup components to prevent failure considering failure probabilities calculated based on incoming traffic. Nonetheless, these models consider a single jurisdiction affected by the disaster.

Coordination Across Jurisdictions

Disasters are often connected across different jurisdictions, with the outcome of one jurisdiction affecting the situation on another one (Ansell et al., 2010) and requiring to manage conflicting objectives (Rui et al., 2008). In these environments, deciding whose resources to use is a complex decision because of the inability to ensure fa ir allocation of joint operations (Schulz & Blecken, 2010), leading to concerns about fair distribution of shared workload (Schulz & Blecken, 2010). The combination of these problems can negatively influence the cost and service level provided.

Research has discussed the problems associated with formal structures in handling new and complex problems (Olsson, 2015) and underlines the importance of adopting cross-sector partnerships in humanitarian logistics (Nurmala et al., 2017). Ansell et al. (2010) argue the need for a transboundary authority structure that can be triggered in times of crisis and that will help minimize confusion and bureaucratic infighting. Collaboration between municipalities and different services, for instance, can enable more effective operations (Zikeloglou et

al., 2024). That is the reason the European Union has led the first European Disaster Response Exercise (Hollis, 2020) and some articles have considered the collaboration between different agencies using agent -based simulation (Lebcir & Roy, 2023; Valaei Sharif et al., 2023), game theory (Coskun et al., 2019; Ergün et al., 2023), and Industry 4.0 technologies (Kumar & Singh, 2022). Nonetheless, less effort has been invested into humanitarian logistics activities. Wojtalewicz et al. (2014) use the case of the Regional Hub Reception Center (RHRC) led by the Purdue Homeland Security Institute involving 17 jurisdictions to propose the use of simulation to efficiently assess the use of facilities to improve coordinated responses. Polese et al. (2024) propose an approach for cross-border multi-risk assessment as an answer for single-risk and independent assessments which neglect the cascading effects between jurisdictions. Wang et al. (2022) use a regional healthcare coalition to collaborate for stock pre-positioning and post-disaster procurement and transshipment to serve affected areas. However, current models are ill-equipped to consider the coordination of different jurisdictions for shelter management.

This paper designs and tests a novel optimization model to allow for the location and allocation of primary and backup shelters enabling the coordination and use of resources from different jurisdictions. The contributions of this article are threefold; i) provides a model allowing operational collaboration between stakeholders from different jurisdictions for shelter management, which is an area neglected in the literature; ii) proposes a capacitated formulation increasing the reliability of the response by acknowledging the possibility of failure of any facility, an approach understudied in the literature; and iii) shows the impact of transboundary coordination on the experience of the people affected by the disaster, which can foster the development of more contributions incorporating transboundary coordination.

METHODOLOGY

This study develops a two-stage optimization model to guide the selection of primary and backup shelters for people affected by disasters considering transboundary coordination. The aim is to find the best combination of the decisions that can lead to a good level of service for disaster beneficiaries.

Optimization is a commonly used method to determine the 'optimal' location of facilities because it allows investigating the different possibilities and define the most appropriate option based on the objective functions (Kınay et al., 2018). It has been selected for this research because it allows balancing the support provided to disaster victims and the resources available during disasters. There have been discussions in the literature about the best performance measures for humanitarian logistics. Although minimization of cost is often used to ensure efficiency in the operations, this research argues that the effectiveness of the operations (i.e. the support given to affected people) needs to be a priority. This research incorporates both objectives in a multi-objective formulation to guide decision-making. The use of two-stage optimization allows us to incorporate the uncertainty about the impact of the disaster and the state of the infrastructure (Rodríguez-Espíndola et al., 2020), where the first sta ge is at the imminent pre-disaster stage and the response stage uses information about the impact of the disaster to guide response decisions.

Context of the Situation

The model proposed promotes the coordination between different jurisdictions to make allocation decisions of primary and backup shelters. The allocation of people is guided by two objective functions. The first objective function minimizes the total distance travelled by the people affected to focus on effectiveness. With the possibility of transboundary coordination, people evacuated in a demand area can move to a different jurisdiction if it is beneficialfor them. To ensure that there is a level of fairness between authorities from different jurisdictions, especially looking at the expenses incurred, the second objective function of the model minimizes the maximum cost per person per jurisdiction. It balances the expected cost (i.e. the sum of facility cost and transportation cost) to avoid overburdening any one of the jurisdictions.

In the two-stage formulation, first stage decisions are focused on facility selection and the flow of people from demand areas to shelters right before the event. Given the range of time for the arrival of people, it is assumed people will get there by their own means. Second-stage decisions take place at the response stage. Considering uncertainty on the magnitude and impact of the disaster, a set of disaster scenarios can be used to guide decision-making. The model incorporates damage to shelters and to the road network to plan the transportation of people allocated to those facilities in buses provided by authorities. Hence, the second stage is focused on the decisions about transportation of people, defining the number of people to move to the same or different jurisdictions.

Notation

Sets

I	Set of demand areas
J	Set of potential shelter locations
L	Set of jurisdictions
D^l	Set of demand areas in jurisdiction l
L^l	Set of shelters in jurisdiction l
S	Set of scenarios

Parameters

TM	Capacity of the buses
TC	Transportation cost per km
$N1$	Percentage of population served by their own jurisdiction at stage 1
$Budget_l$	Operational budget at jurisdiction l
d_{ij}	Distance between demand area i and shelter j
e_{jk}	Distance between shelter j and shelter k
π_s	Probability of scenario s
Cap_j	Capacity of shelter j
Pop_i	Demand in area i
Bus_l	Number of buses available in jurisdiction l
G_j	Cost of locating facility j
H_{ks}	Availability shelter k in scenario s.
W_{jks}	Road connectivity between shelter j and shelter k in scenario s

Decision variables

x_{ij}	Number of people sent from demand area i to shelter j
y_{jks}	Number of people sent from shelter j to shelter k in scenario s
A_j	Use of shelter j
R_{jks}	Number of buses going from shelter j to shelter k in scenario s
M	Maximum operational cost per person

Model

$$Min\ z = \sum_{i\in I}\sum_{j\in J} d_{ij}\, x_{ij} + \sum_{j,k\in J: j\neq k}\ \sum_{s\in S} \pi_s\, e_{jk}\, y_{jks} \tag{1}$$

$$Min\ M \tag{2}$$

$$M \sum_{i\in D^l} Pop_i \geq \sum_{j\in L^l} G_j A_j + \sum_{j\in L^l k\in J: j\neq k}\ \sum_{s\in S} \pi_s TC\, e_{jk}\, R_{jks} \qquad \forall\ l \in L \tag{3}$$

$$\sum_{i\in I} x_{ij} + \sum_{k\in J: j\neq k} (1 - H_{ks}) y_{kjs} \leq Cap_j\, A_j \qquad \forall\ j \in J, s \in S, H_{js} = 1 \tag{4}$$

$$\sum_{j\in J} x_{ij} = Pop_i \qquad \forall\ i \in I \tag{5}$$

$$\sum_{i\in I} x_{ij} = \sum_{k\in J, j\neq k, H_{ks}=1} W_{jks} y_{jks} \qquad \forall\ j \in J, s \in S, H_{js} = 0 \tag{6}$$

$$y_{jks} \leq TM\, R_{jks} \qquad \forall\ j \in J, s \in S, H_{js} = 0 \tag{7}$$

$$\sum_{k\in J, j\neq k} R_{jks} \leq Bus_l \qquad \forall\ l \in L, j \in L^l, s \in S, H_{js} = 0 \tag{8}$$

$$y_{kjs} \leq Cap_k \qquad \forall\ j, k \in J, s \in S, H_{js} = 1, H_{ks} = 0 \tag{9}$$

$$\sum_{i\in D^l}\sum_{j\in L^l} x_{ij} \geq N1 \sum_{i\in D^l} Pop_i \qquad \forall\ l \in L \tag{10}$$

$$A_j \in \{0,1\};\ \ x_{ij}, y_{jks}, R_{jks} \in Z^+$$

Objective function (1) minimizes the total sum of the distance travelled to go to primary and backup shelters. Objective function (2) minimizes the maximum cost of the operations (M) per person per jurisdiction to balance the burden of the different jurisdictions. The cost is calculated in inequality (3) to incorporate the cost of opening facilities and transporting people from primary to backup shelters by the number of people originally residing in the area. Inequality (4) ensures the capacity of the shelters is not exceeded, whereas equation (5) forces the system to allocate all the people in the affected area to a shelter. Equation (6) defines the number of people allocated to a primary shelter that has been damaged to transfer them to the backup shelter. Inequality (7) defines the number of buses needed whilst inequality (8) ensures the number of buses allocated does not exceed the number of available vehicles. Inequality (9) sets a limit on the number of people who can move from a damaged shelter to the backup shelter Inequality (10) is activated if there is an agreement to ensure a percentage of the population of a jurisdiction is served by the same jurisdiction at the first stage tuning the parameter $N1$. Finally, variables are declared.

Solution Method

This research assumes that the overarching priority for authorities is to protect people. Considering the relative importance of objective function (1) over function (2), the model is solved using a lexicographic approach. The model is initially solved to minimize objective function (1) and using objective function (2) as a constraint. Then, the term LEX is introduced to fix the value of the total distance using the following equation:

$$LEX = DIS$$

A new inequality can be defined using the term LEX:

$$LEX \geq \sum_{i \in I} \sum_{j \in J} d_{ij}\, x_{ij} + \sum_{j,k \in J: j \neq k} \sum_{s \in S} \pi_s\, e_{jk}\, y_{jks} \qquad (11)$$

Equation (11) substitutes Equation (1) in the original formulation and it becomes a constraint. The model is then solved minimizing (2) to obtain a single value that can be presented to the decision-maker.

ANALYSIS OF RESULTS

This study is divided into two parts: one using numerical examples and the second one using information from a real case study. This study uses a set of numerical experiments in the first part to analyze the performance of the model, the importance of backup shelters and to explore the value of the stochastic solution in instances of different sizes. The second part is focused on the case of an eruption of the Popocatepetl volcano in Mexico to look at the impact of transboundary coordination under real conditions.

Numerical Instances

Model Results

The model was initially tested using a set of randomly generated numerical examples. The experiments were set based on three instances of different impact shown on Table 1.

Table 1. Characteristics of the numerical examples

Instance	Demand areas p/jurisdiction	Shelters p/jurisdiction	Jurisdictions	Scenarios	Demand
Low	50	50	3	5	51,310
Medium	75	75	4	15	68,991
High	100	100	5	30	86,152

Parameters for the numerical examples were generated using a random number generator based on the type of variables (e.g. binary, integer). The information was coded into GAMS 35.2.0 and solved using Cplex following a lexicographic approach. Each instance was solved allowing a maximum of 12 hours for solution time per model. The process followed was to minimize objective function (1) first, then fixing the objective value, and then minimizing objective function (2) to obtain the solution. A summary of the f indings is presented in Table 2. The optimality gap is defined as the distance between the objective value of the feasible solution found and its optimal value. The first two instances had a gap below 10% within the time limit of 43,200 seconds, whereas the gap of over 15% of the high instance is deemed acceptable because of its size. The results suggest that longer solution

times or heuristic algorithms might be required to close the gap in large-scale instances.

Table 2. Summary of the findings

Inst.	Total distance p/person (km)	Max buses p/1000 people	% of people across jurisdictions (Stage 1)	Max % of people across jurisdictions (Stage 2)	Max % people transferred to backup	Gap (%)	Solution time (s)
L	1.55	0.61	35.54	1.39	2.57	0.18	5.10
M	2.4	3.33	59.84	11.46	13.18	5.6	43,200
H	3.29	4.59	77.86	19.39	21.26	15.85	43,200

The results show a large percentage of people travelling across jurisdictions and requiring backup shelters. Additionally, the model was able to split the expenditure across the different jurisdictions successfully as shown in Table 3. That is also reflected in the number of shelters opened in different jurisdictions, trying to find the best combination of resources without burdening any jurisdiction.

Table 3. Analysis per jurisdiction

Instance	Cost/person/jurisdiction (MXN)		Shelters p/1000 people	
L	R1	62.06	R1	0.31
	R2	62.35	R2	0.35
	R3	61.21	R3	0.31
	R4	-	R4	-
	R5	-	R5	-
M	R1	71.35	R1	0.41
	R2	72.17	R2	0.48
	R3	71.75	R3	0.41
	R4	72.08	R4	0.42
	R5	-	R5	-
H	R1	73.89	R1	0.45
	R2	73.79	R2	0.45
	R3	73.82	R3	0.48
	R4	74.18	R4	0.45
	R5	73.87	R5	0.42

Value of the Stochastic Solutions

To investigate the value of the stochastic solution, we calculated two traditional measures: the expected value of perfect information (EVPI) and the value of the stochastic solution (VSS) shown on Table 4. EVPI is calculated using the value of the recourse problem (RP) and the wait-and-see solution (WS) using the formula: $EVPI = z^{RP} - z^{WS}$. VSS is calculated using the values of the expected value problem (EV) and the Expectation of the Expected Value Solution (EEV) with the formula: $VSS = z^{EEV} - z^{RP}$.

Table 4. Value of the stochastic solution

	EVPI		VSS (Relaxed A and C)	
Instance	Total distance travelled p/person (km)	Max Cost/person /jurisdiction (MXN)	Total distance travelled p/person (km)	Max Cost/person /jurisdiction (MXN)
Low	0.02	12.01	0.25	13.15
Medium	0.14	21.17	0.32	0.82
High	0.29	19.35	0.28	9.13

Figure 1 shows the different solutions obtained for the problem. For all iterations we found that the values obtained

in the RP solution were higher than the values from the wait-and-see problem ($EVPI \in R^+$). The high value of EVPI, particularly in the distance travelled, shows a significant difference between the RP and WS values, which means that the variability in the problem has a significant effect on the number of people served.

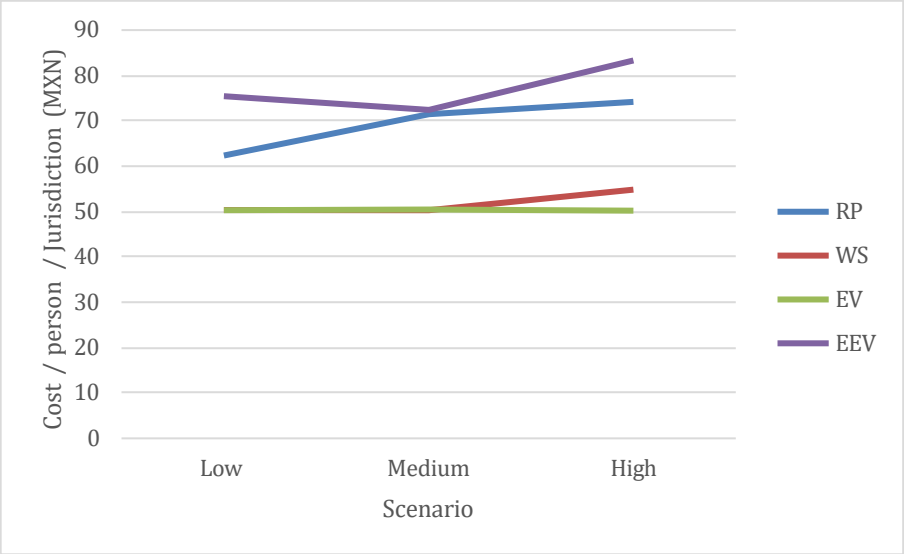


Figure 1. Results of the VSS and EVPI analysis.

The second part of the analysis looked at the cost of ignoring the uncertainty in decision-making. Initially, all the EEV problems were unfeasible when we fixed the variables from the EV solutions. That means that making decisions based on a unique expected value significantly increases the impact on people affected by the disaster. An alternative to obtain feasible EEV values is relaxing some of the first stage variables. After different tests, variable A_j was unfixed to obtain the values shown in the Table. From the practical perspective this result makes sense, given that the availability and capacity of the shelters affect the decisions about backup shelters.

Case Study

Context

The study used real information from a case study in México to investigate the impact of cross-jurisdictional coordination in disaster management. The characteristics of the case can be found in Table 5.

Table 5. Characteristics of the case study

Jurisdiction	Demand areas	Shelters	Demand (people)
EDOMEX	48	60	28,303
PUEBLA	168	141	25,495
MORELOS	107	8	4,998

The eruption of the Popocatépetl volcano in Mexico is used as a case study because of the potential impact on different States of the country. The volcano is located inside a very populated region close to Mexico city, with over 165,000 inhabitants living within 20 km from the crater and surrounded by critical infrastructure (García et al., 2023). It represents one of the latent hazards in the country (medium to high level alert) and it is surrounded by three different States: Estado de Mexico (EDOMEX) Puebla and Morelos, each one of them with response plans. The combination of latent danger of an eruption, a densely populated area and the presence of different jurisdictions make the case relevant to show the application of the model and to draw useful insights.

The purpose of this analysis was to look closer into response mechanisms considering the collaboration (or lack thereof) between the authorities of these States.

Data Collection and Gathering

The demand areas and people to evacuate were estimated using a radius of 30 km of the volcano based on

estimations from Mexican authorities. We used the census data (https://www.inegi.org.mx/programas/ccpv/2020/) and road networks (https://www.gob.mx/imt/acciones-y-programas/red-nacional-de-caminos) from Mexican authorities. The catalogues of facilities with their capacities were obtained from to the States of Mexico (https://cgproteccioncivil.edomex.gob.mx/refugios_temporales), Puebla (https://proteccioncivil.puebla.gob.mx/refugios-temporales) and Morelos (https://obraspublicas.morelos.gob.mx/pdf/rutas-de-evacuacion-y-refugios-temporales). Shelter costs were estimated using the capacities of the shelters and the requirements per person from Mexican authorities (https://www.dof.gob.mx/nota_detalle.php?codigo=5257322&fecha=03/07/2012#gsc.tab=0). The number of buses was assumed to be 1,250 per jurisdiction.

The digitized information was used to determine demand based on distances to the volcano. The distances were calculated using Transcad© after digitizing the facilities found in catalogues from authorities and using the centroids of each one of the AGEBS (Área Geoestadística Básica in Spanish – Basic Geostatistical Area). AGEBs are population units used in Mexico (similar to neighborhoods). Those distances were used to estimate transportation cost based on the consumption of diesel from Serrano-Guevara et al. (2022) and average Diesel cost in Mexico.

Two dimensions were used to generate scenarios: magnitude of the disaster and damage to the infrastructure. Using 4 levels (low, low-medium, medium-high, high), a total of $2^4 = 16$ scenarios were generated for the analysis. The probability of scenarios was generated assuming probabilities of 0.5 for low magnitude/damage, 0.25 for low-medium magnitude/damage, 0.15 for low-high magnitude/damage, and 0.1 for high magnitude/damage. The availability of facilities was defined using a random number generator [0,1].

Findings

The first part of the analysis was concentrated on coordination. The model proposed was compared with two similar models: one not allowing coordination (i.e., each State handles their own population) and another one with partial coordination (i.e., people is directed to shelters in their jurisdiction in the first instance but allowed to go to other jurisdictions for backup shelters). The results are shown in Table 6. Very small gaps show that the solutions obtained are very close to the optimal value in all the instances with solution times below 4 minutes in all cases.

Table 6. Comparison of coordination scenarios

	Total distance travelled (km)	Max cost/person/jurisdiction (MXN)	Gap	Time
Total coordination	1,298,902.38	1,828.76	0.01	204.04
Partial coordination	2,519,124.89	1,096.03	0.01	200.13
No coordination	2,547,299.96	769.18	0.00	180.74

The distance travelled by the affected people is significantly reduced by the mobility across jurisdictions, cutting the value nearly in half. An interesting aspect though is that the cost per person per jurisdiction is considerably lower when no coordination is permitted. To delve closer into that result, Table 7 is used to contrast the cost per person at each one of the jurisdictions, showing that a flow of people travels from EDOMEX and PUEBLA to MORELOS when coordination is allowed.

Table 7. Cost per person per jurisdiction

Cost/person/jurisdiction	EDOMEX	PUEBLA	MORELOS
Total coordination	715.28	891.11	1,828.76
Partial coordination	740.43	872.28	1,096.03
No coordination	713.93	659.92	769.18

We used a sensitivity analysis with partial coordination to see the behavior of the system. The results can be seen in Table 8. All the gaps were below 5%, which shows that the solutions obtained are very close to the optimal value in all the instances. Solution times are below one hour which is considered acceptable for the planning horizon of the model. The model is forcing a percentage of people (from 10% to 100%) to be served in their original State in the first stage, although there is flexibility to move to another jurisdiction in the second stage.

The results show consistent changes in distance and cost when the percentage of people is modified.

Table 8. Sensitivity analysis

Same jurisdiction	Total distance travelled (km)	Max cost/person/jurisdiction (MXN)	Time (s)	Gap
10%	1,320,293.53	1,790.93	253.60	0.17%
20%	1,364,022.34	1,753.10	464.71	4.60%
30%	1,417,004.34	1,734.19	711.82	1.40%
40%	1,490,484.15	1,647.18	962.45	0.69%
50%	1,586,394.40	1,584.34	1,218.67	0.00%
60%	1,710,150.91	1,722.72	1,481.07	0.00%
70%	1,860,978.08	1,402.67	1,736.04	2.80%
80%	2,050,855.03	1,311.74	1,991.38	4.20%
90%	2,280,056.95	1,152.80	2,239.88	0.14%
100%	2,519,124.89	1,096.04	2,478.48	0.91%

DISCUSSION

Numerical Examples

The number of people crossing boundaries to find safe places at stage one is noteworthy. The results show that nearly more than a third of the population evacuated is moving to a shelter in a neighboring jurisdiction in all three instances, with more than half of the population doing so in the medium and large instances. The number increases with the magnitude of the disaster, showing that coordination and collaboration among jurisdictions can be instrumental to protect affected people, especially for large-scale disasters.

Several articles in the literature neglect considering shelter failure and its impact on the population. Our results underscore that backup shelters are particularly important for the medium and high impact instances, although less common in the small instance. These results highlight the importance of backup shelters and transboundary coordination in disaster settings. The findings underscore the potential of the model proposed to balance the burden for the jurisdictions involved. The comparison of the dif ferent instances shows that the model mainta ined a similar level of expenses for all the jurisdictions involved, trying to push down the maximum cost per person per jurisdiction following the second objective function.

The analysis about the value of the stochastic solution derived large values of EVPI and VSS, showing the relevance of the two-stage stochastic formulation. Inadvertently, the unfeasibility of EEV points out a problem already identified in practice. When backup shelters are neglected, there is vulnerability in the system to have a shortage of capacity to shelter all the population needed. The VSS results confirm that neglecting to consider the uncertainty in the different instances has a significant negative effect on the ou tcome of the operations.

Case Study

The reason behind the lower distance traveled when cross-jurisdictional coordination is allowed compared to the instance with no coordination is because different communities from the three States are in close proximity, making it easier for them to move to some of them rather than going across their State.

It was noteworthy to find that the cost per jurisdiction was lower in the instance with no coordination compared to the instance allowing coordination. The closeness between some communities from different States and the lexicographic approach prioritizing distance travelled influence this result. When coordination is allowed people from EDOMEX and PUEBLA go to MORELOS, increasing the overall cost in that State. Once the minimum distance is fixed, the model is unable to find an answer that can balance the cost of the three States. That is an important consequence to have in mind given the importance of reducing the total distance travelled.

The sensitivity analysis shows how, the least flexibility is allowed for people to go to other States, the higher the distance they need to travel to reach a safe destination. That stresses the importance of coordination between different jurisdictions to improve the conditions faced by people affected by disasters. Maximum cost per person per jurisdiction is less clear because it depends on the results of the minimization of the total distance travelled

and the particular network in study.

Overall, the results from the numerical examples and the case study provide three main points: i) coordination and interoperability makes a significant difference in the service provided to people affected by disasters; ii) it is important to contemplate the possibility of failures in primary shelters and the need of secondary/backup shelters to successfully provide housing for people displaced by the disaster; and iii) ignoring the uncertainty in the event when making sheltering decisions has a significant negative effect in the performance of humanitarian operations.

CONCLUSIONS

This article introduces a two-stage stochastic model for the selection of primary and backup shelters for vulnerable population considering the collaboration between different jurisdictions. This study is the first one in the literature to incorporate transboundary coordination into primary and backup shelter location. The literature on transboundary coordina tion is still at a nascent stage, but the findings from this study show the impact of ignoring the interaction between authorities from different boundaries. Although jurisdictional and political issues can affect disaster planning, this article provides evidence that a well-coordinated joint response can be beneficial for the victims, and even for the authorities involved.

There are different opportunities for further research. The extension of policies such as the p -next center problem (Albareda-Sambola et al., 2015) can be valuable to advance similar models in the future. The experimentation with different objective functions could be useful to see the behavior of the model under different policies. Finally, the development of a heuristic to reduce solution times for larger instances would be beneficial for the application of the model in practice.

REFERENCES

Aghaie, S., & Karimi, B. (2022). Location-allocation-routing for emergency shelters based on geographical information system (ArcGIS) by NSGA-II (case study: Earthquake occurrence in Tehran (District-1)) [Article]. *Socio-Economic Planning Sciences*, *84*, Article 101420. https://doi.org/10.1016/j.seps.2022.101420

Albareda-Sambola, M., Hinojosa, Y., Marin, A., & Puerto, J. (2015). When centers can fail: A close second opportunity. *Computers & Operations Research*, *62*, 145-156. https://doi.org/10.1016/j.cor.2015.01.002

Albareda-Sambola, M., Martínez-Merino, L. I., & Rodríguez-Chía, A. M. (2019). The stratified p-center problem. *Computers & Operations Research*, *108*, 213-225. https://doi.org/https://doi.org/10.1016/j.cor.2019.04.013

An, S., Cui, N., Li, X., & Ouyang, Y. (2013). Location planning for transit-based evacuation under the risk of service disruptions. *Transportation Research Part B: Methodological*, *54*, 1-16. https://doi.org/https://doi.org/10.1016/j.trb.2013.03.002

Ansell, C., Boin, A., & Keller, A. (2010). Managing Transboundary Crises: Identifying the Building Blocks of an Effective Response System. *Journal of Contingencies and Crisis Management*, *18*(4), 195-207. https://doi.org/10.1111/j.1468-5973.2010.00620.x

Babashahi, M., Shahanaghi, K., Gholamian, M. R., & Yavari, A. (2020). A bi-objective hierarchical hub location model with facility failure. *Scientia Iranica*, *27*(4), 2009-2020. https://doi.org/10.24200/sci.2018.5358.1229

Balcik, B., Beamon, B. M., Krejci, C. C., Muramatsu, K. M., & Ramirez, M. (2010). Coordination in humanitarian relief chains: Practices, challenges and opportunities [Article]. *International Journal of Production Economics*, *126*(1), 22-34. https://doi.org/10.1016/j.ijpe.2009.09.008

Boin, A., Rhinard, M., & Ekengren, M. (2014). Managing Transboundary Crises: The Emergence of European Union Capacity. *22*(3), 131-142. https://doi.org/https://doi.org/10.1111/1468 -5973.12052

Coskun, A., Elmaghraby, W., Karaman, M. M., & Salman, F. S. (2019). Relief aid stocking decisions under bilateral agency cooperation. *Socio-Economic Planning Sciences*, *67*, 147-165. https://doi.org/https://doi.org/10.1016/j.seps.2018.10.009

CRED. (2023). *2022 Disasters in Numbers*. Retrieved 8th August from https://www.preventionweb.net/publication/2022-disasters-numbers?utm_source=PreventionWeb&utm_campaign=8d268f2c6f - PreventionWeb+daily&utm_medium=email&utm_term=0_b73053c1c6 -8d268f2c6f-485859101

Ergün, S., Usta, P., Alparslan Gök, S. Z., & Weber, G. W. (2023). A game theoretical approach to emergency logistics planning in natural disasters. *Annals of Operations Research*, *324*(1), 855-868. https://doi.org/10.1007/s10479-021-04099-9

Gangwal, U., Siders, A. R., Horney, J., Michael, H. A., & Dong, S. (2023). Critical facility accessibility and road criticality assessment considering flood-induced partial failure [Article]. *Sustainable and Resilient Infrastructure*, *8*(sup1), 337-355. https://doi.org/10.1080/23789689.2022.2149184

García, A. R., Zavala-Hidalgo, J., Delgado-Granados, H., Garcia-Escalante, J., Gómez-Ramos, O., & Herrera-Moro, D. (2023). An automated ash dispersion forecast system: case study Popocatépetl volcano, Mexico. *Journal of Applied Volcanology*, *12*(1), 9. https://doi.org/10.1186/s13617-023-00135-4

Geng, S., Hou, H., & Zhou, Z. (2021). A hybrid approach of vikor and bi-objective decision model for emergency shelter location–allocation to respond to earthquakes [Article]. *Mathematics*, *9*(16), Article 1897. https://doi.org/10.3390/math9161897

Gutjahr, W. J., Noyan, N., Vandaele, N., & Van Wassenhove, L. N. (2020). Innovative approaches in humanitarian operations. *OR Spectrum*, *42*(3), 585-589. https://doi.org/10.1007/s00291-020-00598-6

Hinojosa, Y., Marín, A., & Puerto, J. (2023). Dynamically second-preferred p-center problem. *European Journal of Operational Research*, *307*(1), 33-47. https://doi.org/https://doi.org/10.1016/j.ejor.2022.09.019

Hollis, S. (2020). Crisis management in Europe: exploring transgovernmental solutions to transboundary problems [Article]. *Journal of Transatlantic Studies*, *18*(2), 231-252. https://doi.org/10.1057/s42738-020-00042-1

Kınay, Ö. B., Yetis Kara, B., Saldanha -da-Gama, F., & Correia, I. (2018). Modeling the shelter site location problem using chance constraints: A case study for Istanbul. *European Journal of Operational Research*, *270*(1), 132-145. https://doi.org/https://doi.org/10.1016/j.ejor.2018.03.006

Kumar, P., & Singh, R. K. (2022). Application of Industry 4.0 technologies for effective coordination in humanitarian supply chains: a strategic approach [Article]. *Annals of Operations Research*, *319*(1), 379-411. https://doi.org/10.1007/s10479-020-03898-w

Lebcir, R. M., & Roy, P. (2023). Impact of coordination on post-earthquake last mile relief distribution

operations in India [Article]. *International Journal of Emergency Management*, *18*(3), 293-316. https://doi.org/10.1504/IJEM.2023.132390

Liu, Q., Ruan, X., & Shi, P. (2011). Selection of emergency shelter sites for seismic disasters in mountainous regions: Lessons from the 2008 Wenchuan Ms 8.0 Earthquake, China. *Journal of Asian Earth Sciences*, *40*(4), 926-934. https://doi.org/http://dx.doi.org/10.1016/j.jseaes.2010.07.014

Ma, D., Chu, J., Wang, Z., & Sun, T. (2016). Study on disaster shelter complement location problem after disaster shelter failure [Article]. *World Earthquake Engineering*, *32*(1), 117-123. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84963545904&partnerID=40&md5=9610266f5bb4fa197ea4b4d19d915954

Mohamadi, A., Yaghoubi, S., & Pishvaee, M. S. (2019). Fuzzy multi-objective stochastic programming model for disaster relief logistics considering telecommunication infrastructures: a case study. *Operational Research*, *19*(1), 59-99. https://doi.org/10.1007/s12351-016-0285-2

Nurmala, N., de Leeuw, S., & Dullaert, W. (2017). Humanitarian-business partnerships in managing humanitarian logistics. *Supply Chain Management*, *22*(1), 82-94. https://doi.org/http://dx.doi.org/10.1108/SCM-07-2016-0262

Oloruntoba, R. (2010). An analysis of the Cyclone Larry emergency relief chain: Some key success factors. *International Journal of Production Economics*, *126*(1), 85-101. https://doi.org/https://doi.org/10.1016/j.ijpe.2009.10.013

Olsson, E. (2015). Transboundary crisis networks: The challenge of coordination in the face of global threats. *Risk Management*, *17*(2), 91-108. http://www.jstor.org/stable/43695458

Polese, M., Tocchi, G., Babič, A., Dolšek, M., Faravelli, M., Quaroni, D., . . . Prota, A. (2024). Multi-risk assessment in transboundary areas: A framework for harmonized evaluation considering seismic and flood risks [Article]. *International Journal of Disaster Risk Reduction*, *101*, Article 104275. https://doi.org/10.1016/j.ijdrr.2024.104275

Rahmani, D., Zandi, A., Peyghaleh, E., & Siamakmanesh, N. (2018). A robust model for a humanitarian relief network with backup covering under disruptions: A real world application. *International Journal of Disaster Risk Reduction*, *28*, 56-68. https://doi.org/https://doi.org/10.1016/j.ijdrr.2018.02.021

Rodríguez-Espíndola, O., Alem, D., & Pelegrin Da Silva, L. (2020). A shortage risk mitigation model for multi-agency coordination in logistics planning. *Computers & Industrial Engineering*, *148*, 106676. https://doi.org/https://doi.org/10.1016/j.cie.2020.106676

Rui, C., Sharman, R., Rao, H. R., & Upadhyaya, S. J. (2008). COORDINATION IN EMERGENCY RESPONSE MANAGEMENT [Article]. *Communications of the ACM*, *51*(5), 66-73. https://doi.org/10.1145/1342327.1342340

Schulz, S. F., & Blecken, A. (2010). Horizontal cooperation in disaster relief logistics: Benefits and impediments [Article]. *International Journal of Physical Distribution and Logistics Management*, *40*(8), 636-656. https://doi.org/10.1108/09600031011079300

Serrano-Guevara, O. S., Huertas, J. I., Mogro, A. E., & Quirama, L. F. (2022). Consumo energético de vehículos pesados en México. *Informador Técnico*, *87*(1), 29-39. https://doi.org/10.23850/22565035.5123

Valaei Sharif, S., Habibi Moshfegh, P., & Kashani, H. (2023). Simulation modeling of operation and coordination of agencies involved in post-disaster response and recovery [Article]. *Reliability Engineering and System Safety*, *235*, Article 109219. https://doi.org/10.1016/j.ress.2023.109219

Wang, Q., Liu, Z., Jiang, P., & Luo, L. (2022). A stochastic programming model for emergency supplies pre-positioning, transshipment and procurement in a regional healthcare coalition. *Socio-Economic Planning Sciences*, *82*, 101279. https://doi.org/https://doi.org/10.1016/j.seps.2022.101279

Wang, Y., & Xu, Z. (2023). An Emergency Shelter Location Model Based on the Sense of Security and the Reliability Level. *Journal of Systems Science and Systems Engineering*, *32*(1), 100-127. https://doi.org/10.1007/s11518-023-5550-3

Wojtalewicz, C., Kirby, A., & Dietz, J. E. (2014). Implementation and modeling of a Regional Hub Reception Center during mass evacuation operations. *J Emerg Manag*, *12*(3), 197-210. https://doi.org/10.5055/jem.2014.0172

Yahyaei, M., & Bozorgi-Amiri, A. (2019). Robust reliable humanitarian relief network design: an integration of shelter and supply facility location. *Annals of Operations Research*, *283*(1), 897-916. https://doi.org/10.1007/s10479-018-2758-6

Zikeloglou, I., Lekkas, E., Lozios, S., & Stavropoulou, M. (2024). Community's evacuation planning and response for the 2021–2022 fire seasons in Greece. *Safety Science*, *172*, 106434. https://doi.org/https://doi.org/10.1016/j.ssci.2024.106434

Using Optimization Modeling to Determine Locations for Aerial Firefighting Resources Based on Wildfire Risk Data

Viktor Sköld Gustafsson
Linköping University
Department of Science and Technology
viktor.skold.gustafsson@liu.se

Tobias Andersson Granberg
Linköping University
Department of Science and Technology
tobias.andersson.granberg@liu.se

Martin Waldemarsson
Linköping University
Department of Science and Technology
martin.waldemarsson@liu.se

ABSTRACT

The increase of the global mean temperature likely leads to more frequent and prolonged fire weather periods, with hot and dry conditions. This would also lead to increased frequency and magnitude of wildfires, and increased risk of spatial and temporal co-occurrence. A key to control and suppress wildfires during fire weather is a strong initial response, for which aerial firefighting resources can be used. This paper investigates the potential of optimization modeling to support pre-wildfire location planning of aerial firefighting resources, by developing a binary integer programing model and test it on Swedish case data. The model produces useful results given considerably less amount of available input data to determine the locations, compared with the manual planning. With relatively short solution times, it shows potential to support fast relocation of resources during sudden changes of circumstances, and to facilitate implementation of a standard operating procedure, with less expert knowledge dependence.

Keywords

wildfire suppression, initial attack, fire weather, preparedness, planning.

INTRODUCTION

Due to increased amounts of greenhouse gases in the atmosphere, the global mean temperature has increased, and evidence suggest that the warming trend will continue (IPCC, 2021). The mean temperature change will affect temperature extremes, which are already observed in northern Europe (IPCC, 2021), and the warming trend will lead to increased frequencies of all kinds of hot extremes (Seneviratne et al., 2021). Due to strong correlation, increased frequencies of hot extremes will likely lead to more frequent and longer periods of dry conditions or droughts (Zscheischler & Seneviratne, 2017). For the reasons above, there will be more periods with fire weather, which is related to hot and dry conditions, and driven by high temperatures, and low soil moisture and humidity (Turco et al., 2017). More and longer periods of fire weather, will likely lead to increased frequency and magnitude of wildfires (Mueller et al., 2020), with increased risk that wildfires occur spatially and temporally close to each other. This increased risk will require improved preparedness for wildfires.

A key to control and suppress wildfires during periods of fire weather is a strong initial attack (Pierce 1970, via Greulich, 2003), which can be done using aerial firefighting resources (e.g., helicopters and scooping airplanes) (Plucinski et al., 2007). The effectiveness of initial attacks using aerial firefighting resources is restricted by the response times (McCarthy, 2003), which depends on the distance between fire ignition points and the resource locations. This calls for research on pre-wildfire locations of aerial firefighting resources as an effort to improve

the preparedness for future events. To date, this is a underdeveloped area of research (Belval et al., 2022).

The purpose of this paper is to investigate if optimization modelling of short-term pre-wildfire location of aerial firefighting resources can improve the preparedness for wildfires. This is done by developing an optimization model to determine where to locate the resources based on the current risk of wildfires. The results of the optimization model are then compared with a real case to determine its potential to enhance the preparedness for wildfires. The real case includes the historical locations of aerial firefighting resources in Sweden, as well as the historical incidents where the resources were used, during the high-risk wildfire season (June-August) of 2021.

Related Work

According to previous research, some factors are considered critical for the success and effectiveness of managing and suppressing wildfires. In a retrospect analysis of wildfires where aerial suppression was used, Plucinski et al. (2007) finds that the area burning on arrival, the fire danger index, the time to initial attack, and the fuel hazard score were all statistically significant factors to the success of the initial attack. In a similar analysis of aerial suppression effectiveness in Australia, McCarthy (2003) also finds the time to initial attack as a critical factor for the effectiveness, and argues that the earlier the resources can reach the fire, the more likely they will be able to contain it.

Regarding the time to initial attack, previous studies have used different quantitative methods to address it, from visualizations of retrospective analyses (Belval et al., 2022) to more advanced optimization models (Shahparvari et al., 2021). For example, Greulich (2003) developed a spread-sheet based procedure to determine locations of air bases to improve the preparedness for a given protection area. For a similar purpose, Fried et al. (2006) used stochastic simulation to, for instance, determine long-term locations for the stationing of firefighting resources. More recent studies chose to use the same method as presented in this paper, which is optimization modeling. Both Griffith et al. (2017) and Shahparvari et al. (2021) have developed models to address resource allocation to ongoing fires. The former developed a model to solve dynamic resource allocation problems to ongoing suppression efforts, based on the fire spread direction and environmental conditions (Griffith et al., 2017). The latter focused on resource coordination during suppression efforts requiring collaboration between aerial and ground firefighting resources, by optimizing the allocation of resources and the assignment of tasks required in the effort (Shahparvari et al., 2021).

An example of a study similar to the one presented in this paper is Maritano et al. (2016), who developed a model to determine a helipad network on Sicily, Italy. However, the network is not primarily used to increase the preparedness for wildfires, but a set of services performed by helicopter (e.g., transport, medical services, tourism services, civil protection). Also, the locations they suggest are long term implementations, and the location of helipads is not based on wildfire risk data. The only study found that is closely related to the study presented in this paper is Zeferino (2020), who developed a model to locate aerial firefighting resources in Portugal. The differences in this paper are that we consider tactical, short-term locations, that we have less resources available to locate, and that our locations are based on wildfire risk data.

One contribution of this paper is an optimization model to determine the best locations of a finite set of aerial firefighting resources, based on wildfire risk data. As far as we know, this kind of model has not been developed and validated before. Another contribution is that the model is tested and compared to the historical, manual planning, which is done by experts. The results show that the model manages to produce results that are close to what the experts achieve, by using only wildfire risk data, while the experts have access to much more information. This proves the usefulness of both the Fire Weather Index (FWI) that we use to estimate wildfire risk, and optimization modeling as a base for decision support. In wildfire suppression practices, the model could potentially be used as a support for day-to-day dynamic resource preparedness planning to increase the effectiveness of aerial firefighting resources used for initial attacks.

Method

In this paper, an operations research approach was used by applying optimization modeling to solve a set of instances to the aerial firefighting preparedness location (AFPL) problem. The input data were processed using the programming language R through the platform RStudio, while the model was formulated using the language AMPL through the platform AMPL IDE. Through AMPL IDE, the model and the specific input data for each instance of the problem were sent to the solver CPLEX. Through a case study, the results of the model were tested and compared with aerial firefighting preparedness locations from a real case.

A MODEL FOR THE AERIAL FIREFIGHTING PREPAREDNESS LOCATION PROBLEM

This section presents the problem studied and the assumptions made in our formulation. Furthermore, the mathematical formulation of the model is presented, including sets, parameters, variables, objective function, and constraints.

Problem Description and Assumptions

The problem to be solved in this paper is called the aerial firefighting preparedness location (AFPL) problem, here solved as a *p-median* problem (Daskin & Maass, 2015). The problem can be stated as

To determine a finite number of preparedness locations for aerial firefighting resources that minimizes the weighted distance between the locations and all risk points based on their wildfire risk value.

The goal is to minimize the total "risk distance", which is the sum of the products between the distance and the risk value. The following assumptions have been made:

- The risk value for each risk point is known
- There is a finite number of resource locations to be determined
- Since the resources are aerial, Euclidian distances between each pair of location and risk point can be used
- Although aerial firefighting resources might have different characteristics in reality (e.g., speed, capacity), the resources in the model all have the same characteristics

Model Formulation

The AFPL problem was formulated as a binary integer program (BIP). The notation used in the mathematical representation of the model is presented below.

Sets

$I = \{1,\ldots, n\}$, *the set over all risk forecast points*
$J = \{1,\ldots,m\}$, *the set over all possible resource location points*

Parameters

d_{ij} = *the Euclidean distance between risk point i and location point j*
h_i = *the risk value in risk forecast point i*
p = *the number of aerial resources to be located*

Variables

$$yy_{jj} = \begin{cases} 1, & \text{if a resource is located at point } jj \\ 0, & \text{otherwise} \end{cases}$$

$$xx_{iijj} = \begin{cases} 1, & \text{if risk point } ii \text{ is assigned to a resource located at point } jj \\ 0, & \text{otherwise} \end{cases}$$

Using the above notation, the objective function (1) and constraints (2)-(5) below were constructed.

Objective Function and Constraints

$$\text{Min} \sum_{ii=1}^{nn} \sum_{jj=1}^{mm} h_{ii} ss_{iijj} xx_{iijj} \tag{1}$$

$$\text{s.t.} \quad yy_{jj} \geq xx_{iijj} \qquad \forall ii \in II, jj \in JJ \tag{2}$$

$$\sum_{jj=1}^{mm} xx_{iijj} \geq 1 \qquad \forall ii \in II \tag{3}$$

$$\sum_{jj=1}^{mm} yy_{jj} = pp \tag{4}$$

$$xx_{iijj}, yy_{jj} \in \{0,1\} \qquad \forall ii \in II, jj \in JJ \tag{5}$$

The model is solved for a specific day and finds operational locations for the resources to match the risk values that day. The objective (1) is to minimize the distance between resource locations and risk points, based on their risk values. Hence, the goal is to minimize the total "risk distance", which is the product between h_i and d_{ij} for all risk points *i* covered by the closest resource location *j*. This corresponds to minimizing the total time spent on travelling to incidents, given that the number of incidents in a risk point, over time, is directly proportional to the risk value. The constraints in (2) ensure that a risk point *i* can only be covered by a selected resource location *j* if there is a resource located there. The constraints in (3) ensure that all risk points *i* are covered by at least one resource location *j*. The single constraint in (4) ensures that the total number of resource locations selected must be equal to a finite number *p*. In this paper, the value of *p* is determined by the historical number of resources located during each studied date.

CASE STUDY

To test and validate the potential of the model, a case study was performed for Sweden using data from the wildfire high-risk season (June-August) of 2021. Looking at days when the aerial resources were used at a wildfire effort, a comparison was made between the locations that were suggested by the model (BIP locations) and the historical locations (Real locations) determined by experts before the effort. For each wildfire location, the distance from the closest resource was calculated and compared.

The Swedish Decision-Making Process

In Sweden, the aerial firefighting preparedness locations are decided by the Directions and Priority Function (IPF), with representatives from different departments of the Swedish Civil Contingencies Agency (MSB). The IPF holds regular meetings twice a week during the fire season, with the possibility of increased frequency during periods with multiple ongoing missions or high fire risk and uncertainty.

Before the meeting, an analysis is performed and documented based on the weather and fire risk forecasts. A fire risk forecast is available on daily level for the coming six days, and on hourly level for the coming 48 hours. Beside the forecast data, the conducted analysis is based on the forecast uncertainty and reports from external actors in the Swedish emergency response system. To complement the nationwide analysis, critical and uncertain geographical areas may be studied further, especially concerning previous weather, extreme fire behavior, and the ability to manage a potential fire in the area.

After an IPF meeting, the preparedness location proposals are put forward to the deputy decision maker of the MSB, and if approved, a formal decision is signed and documented. The decision is formulated as the number of resources (helicopters and planes, respectively) that should be prepared to respond, together with their determined locations during the specified period. The decision is then forwarded as orders to the operational unit level.

Input Data

The main set of input data used covers the utilization of aerial firefighting resources, constituting scooping airplanes and helicopters, in wildfire suppression efforts during the case period. In total, the set consists of 54 suppression efforts where aerial resources were initiated, during 34 different dates. This means that the case included 34 instances of the AFPL problem to be solved. For each date where aerial resources were used, a set of wildfire risk data was provided by the Swedish Meteorological and Hydrological Institute. Each date included a set of 62 351 points with risk values based on the Fire Weather Index (FWI). The FWI is a compiled index accounting for the effects of soil moisture and wind on the fire's behavior in terms of build-up and spread (National Wildfire Coordinating Group, 2021). However, these points were aggregated into a network of 1 335 20x20 km zones, where the FWI value of each point included in each zone were added together.

The set of possible resource locations included 286 points, representing the center coordinates of the most densely populated cities in each municipality of Sweden. The reason for using this set is that the current selection of preparedness locations in Sweden is on municipal level. The number of resource locations is finite in each instance of the problem. Thus, the finite number denoted *p* in the formulation was set to the real number of resource

locations determined. This enables comparison between the model's solution and the locations used in the real case. In summary, each of the 34 case instances of the AFPL problem included:

- 286 possible resource locations
- 1335 risk points
- 1335 FWI values, one for each risk point
- A distance matrix with the distance between each pair of risk point and possible resource location.
- A finite number of selected resource locations (varies between 2 and 8)

COMPUTATIONAL RESULTS

Each of the 34 instances were solved to integer optimality using the solver CPLEX, with a relative MIP gap tolerance of 0.0001. The average solution time over the instances was 764 seconds, with the minimum and maximum time at 217 and 1162 seconds, respectively. The problems were solved on a PC with an Intel® Core™ i7-8565U CPU at 1.8-1.99 GHz, and with a RAM of 16 GB.

To investigate the potential of the developed model, a comparison was made with the real resource locations during the 34 days. The analysis compared the resource locations of the model and the real case over the distance to the 54 efforts where aerial firefighting resources were used. Aggregated results are presented in Table 1.

Table 1. Aggregated results for the BIP model and the real case locations, respectively, over 54 efforts during the high-risk wildfire season of 2021

	BIP Model	Real case
Average distance to effort	96 757 m	91 046 m
Minimum distance to effort	10 206 m	8 631 m
Maximum distance to effort	205 922 m	302 548 m
Occurrences, distance > 200 000 m	1	4
Number of times closest	18	32
Average difference when closest	66 025 m	46 776 m

As can be seen in Table 1, the experts on average manage to locate the resources closer to the efforts. However, looking at efforts where the closest resources are far away (more than 200 km), the model seems to be doing a better job of ensuring that all efforts can be reached within a reasonable time. The experts consider a risk point to be reached within an hour of travelling time if it is within 200 km from an aerial resource, which is why it was used in the comparison.

For the results in Table 1, it is assumed that the resource located closest to each effort point were used. Thus, we assume that if there are two or more efforts during the same date, the location point closest to the efforts will have enough resources to initiate an attack, even if the effort points share the same closest resource location. Since this is not always realistic, further comparative analyses were done for a set of cases with several efforts during the same date. One interesting date is the 7th of August 2021, during which 5 efforts with aerial firefighting resources were initiated. The FWI values, the resource locations, and the effort points for the model and the real case scenario are presented in the two plots in Figure 1.

The assumption that all effort points will use resources from the closest resource location point means that the real historical locations have a better preparedness for the scenario in Figure 1. However, if it instead is assumed that all resource points can only initiate an attack to one effort point, the situation is the opposite. An overview of the results with the respective assumption, using a greedy approach for the model and the real case, for all days with more than one effort initialized, is presented in Table 2. Here, greedy means that for all efforts, the one with the closest available resource gets a resource allocated to it. Then, the effort with the closest available resource, not counting the ones that have already been allocated, gets a resource, and so on.

The change in results when applying the two assumptions (Table 2), comes from the changes of maximal distances for the model and the real case. This change is more tangible for the real case, which means that the suggestion of locations by the model seems more stable during days with several efforts. Only looking at the current fire risk,

the model performs better in terms of being closest to the areas with high wildfire risk. However, the overall results suggest that the current preparedness planning performs better than the model in terms of the distance to efforts where aerial resources were initialized. In the following section, a discussion about possible causes for this will be made.

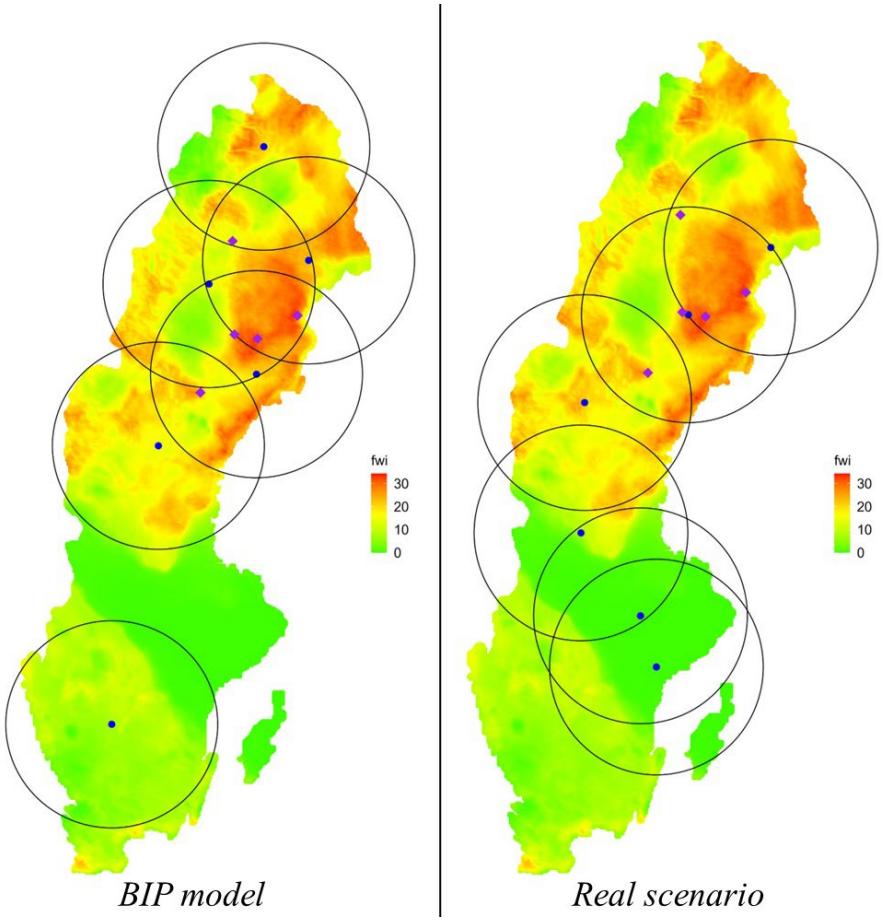


Figure 1. The FWI risk values (green-red), resource location points (blue points), and initiated effort points (purple diamonds) during the 7th of August 2021. The black circles represent the reach (200 km) within one hour from takeoff

Table 2. Aggregated results over the 11 days with more than one effort initialized, with a total of 31 efforts, for the BIP model and the real case scenario, respectively, while influenced by two different assumptions

	Closest resource location		**Closest available resource location**	
	BIP	**Real case**	**BIP**	**Real case**
Average effort distance	90 952 m	83 466 m	122 237 m	128 632 m
Minimal effort distance	10 206 m	8 631 m	10 206 m	8 631 m
Maximal effort distance	193 334 m	257 119 m	481 692 m	747 161 m
Distance > 200 000 m	0	1	4	5
Number of times closest	8	21	9	20

DISCUSSION

The first things to discuss are the situations where the real case locations are close to effort points with low or moderate fire risk. Such cases are extremely difficult, if not impossible, to include in a preparedness planning using optimization modeling, at least without including more aspects than the fire risk in the objective function. Now this is the sole factor determining the resource locations in our model. Figure 2 presents the locations and efforts from the 1st of June 2021, which is such a case.

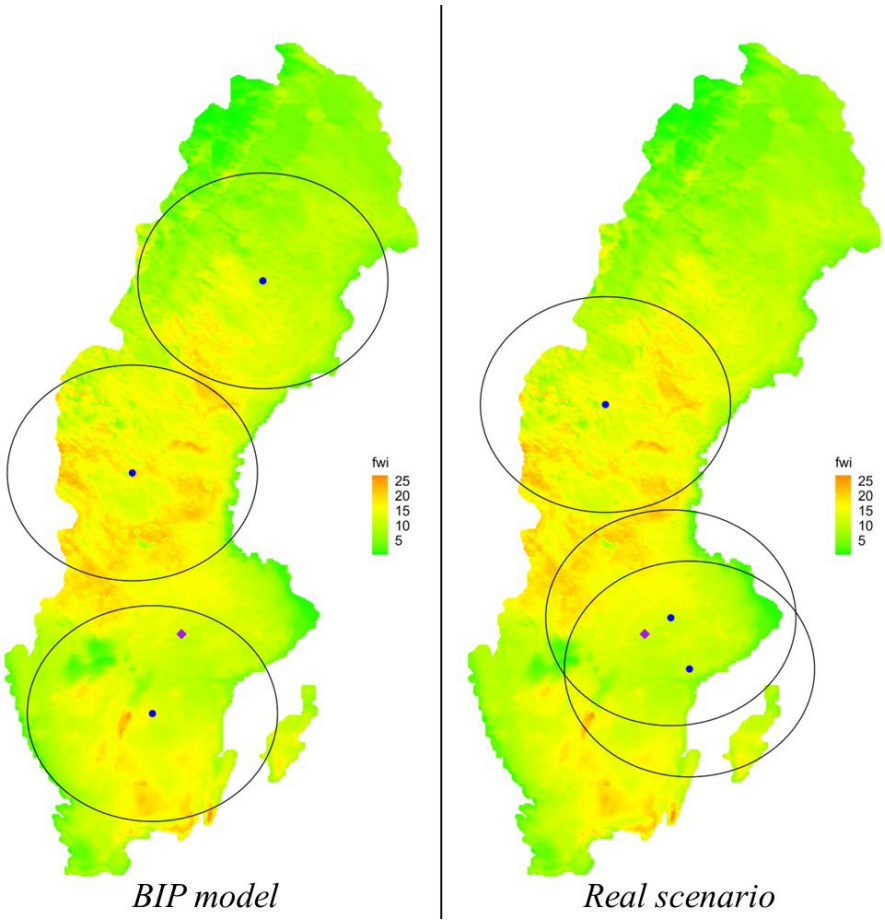


Figure 2. The FWI risk values (green-red), resource location points (blue points), and initiated effort points (purple diamonds) during the 1st of June 2021. The black circles represent the reach (200 km) within one hour from takeoff

In this case, the effort is a wildfire in Arboga, not far from Västerås, which is one of the fixed locations for firefighting helicopters currently used in Sweden. As can be seen in Figure 2, the resource locations are much closer to the initiated effort point in the real case scenario compared to the BIP model. Since the wildfire risk in the effort point is small compared to other areas in Sweden during this time, the BIP model would never suggest locating resources any closer to that point. However, it might be argued that the BIP model determines a more equal resource location regarding the actual fire risk, which is one of the critical factors for initial response success (Plucinski et al., 2007).

Although the comparative results show that the current resource preparedness locations are better on average, we argue that the BIP model shows potential for several reasons. Firstly, it determines the resource locations with considerably less amounts of information compared to the real case. Despite this, the difference in average distance to the effort points is only about 7 km. With information available in the real case, e.g., status of ongoing wildfire efforts, resource availability, and the ability to reach the effort point with ground resources, the BIP model might give even better results. The second reason is the solution time. For the instances of the AFPL problem tested, the solution time for the BIP model was on average 13 minutes (with minimum 4 and maximum 19 minutes). With the intended use as input and decision support during pre-planned IPF meetings, these solution times can be considered adequate. It may also enable optimized relocation of resources if the circumstances suddenly change, which is considerably more difficult, complex, and time-consuming without the model. However, of course there still exist improvement opportunities regarding the solution time. Thirdly, the BIP model

enables the implementation of a standard operating procedure in the location planning of aerial firefighting resources. Due to the complexity of the current planning procedure, it is dependent on the knowledge of a small group of experts. With a support tool including the BIP model, the planning procedure can become more standardized, hence, less dependent on individual knowledge and skill. Lastly, the model shows potential in minimizing the maximal distances to efforts during instances of the AFPL with several simultaneous effort initializations. This can be seen in Table 2, where the BIP model does have longer minimum distances, but considerably shorter maximum distances.

However, to fully investigate the potential of the BIP model as a support tool for resource location planning, more data is needed on the utilization of aerial firefighting resources during wildfire suppression efforts. Although the dataset used includes all aerial fire suppression efforts in the selected period, they count to 54 separate efforts on 34 different dates, which is too few to be certain about the model's potential. Also, to enable a fair and thorough comparison between the model and the real case scenario, their locations need to be based on the same risk data, which currently is not the case. Here, the real case locations are based on forecast data, while the model's locations are based on analyzed risk data. Therefore, this paper should be considered a first step to investigate the potential of optimization modeling for resource preparedness planning, which is identified as a research area with high potential (Belval et al., 2022).

CONCLUSIONS

With increased recognition of the interactions between climate change and the frequency and magnitude of wildfires, more attention is needed on how to improve the preparedness for wildfires in emergency response systems. In a first step to address this, we have investigated the potential of optimization modeling to support pre-wildfire location planning of aerial firefighting resources, utilizing available wildfire risk data based on the Fire Weather Index. This, by developing a BIP model for resource location planning and test it on a case which included 54 separate efforts utilizing aerial fire suppression on 34 different dates during the high-risk fire season 2021 in Sweden. The model's results were then compared with the real resource location planning during this period. Although the model did not outperform the real resource location planning in all cases, it showed potential concerning solution time and the implementation of a standard operating procedure. Additionally, given the considerably less amount of information the model results are based on, the overall results can be argued to be satisfying.

To fully investigate the potential of optimization modelling to support pre-wildfire resource location planning, more research is needed. The future research plan includes incorporating extended wildfire risk data which provides risk forecasts not only for the current day, but a few consecutive days. The plan also includes extensions in terms of other types of data besides the wildfire risk to include in the model, for example resource availability, wildfire reachability, population density, vulnerability, and critical infrastructure.

ACKNOWLEDGEMENTS

The authors would like to thank the Swedish Civil Contingencies Agency and the Swedish Meteorological and Hydrological Institute for providing data and information to this study. The project behind this study, Efficient management of multiple natural events (EMMUNE), is part of the Centre for Advanced Research in Emergency Response (CARER). This work was supported by the Swedish Civil Contingencies Agency and Formas, a Swedish research council for sustainable development [grant 2019-06052].

REFERENCES

Belval, E. J., Short, K. C., Stonesifer, C. S., & Calkin, D. E. (2022). A Historical Perspective to Inform Strategic Planning for 2020 End-of-Year Wildland Fire Response Efforts. *Fire*, *5*(2), 1–18. https://doi.org/10.3390/fire5020035

Daskin, M. S., & Maass, K. L. (2015). The p-median problem. In G. Laporte, S. Nickel, & F. Saldanha de Gama (Eds.), *Location Science*. Springer.

Fried, J. S., Gilless, J. K., & Spero, J. (2006). Analysing initial attack on wildland fires using stochastic simulation. *International Journal of Wildland Fire*, *15*(1), 137–146. https://doi.org/10.1071/WF05027

Greulich, F. E. (2003). Airtanker initial attack: A spreadsheet-based modeling procedure. *Canadian Journal of Forest Research*, *33*(2), 232–242. https://doi.org/10.1139/x02-176

Griffith, J. D., Kochenderfer, M. J., Moss, R. J., Mišić, V. V, Gupta, V., & Bertsimas, D. (2017). Automated Dynamic Resource Allocation for Wildfire Suppression. *Lincoln Laboratory Journal*, *22*(2), 38–59.

IPCC. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC.

Maritano, L., Amoroso, S., & Castelluccio, F. (2016). Heliport network planning through or methods and use of GIS. *Aircraft Engineering and Aerospace Technology*, *88*(3), 365–373. https://doi.org/10.1108/AEAT-01-2015-0016

McCarthy, G. J. (2003). *Effectiveness of aircraft operations by the Department of Natural Resources and Environment and The Country Fire Authority 1997-98*. *Research Report No. 52*, 32.

Mueller, S. E., Thode, A. E., Margolis, E. Q., Yocom, L. L., Young, J. D., & Iniguez, J. M. (2020). Climate relationships with increasing wildfire in the southwestern US from 1984 to 2015. *Forest Ecology and Management*, *460*(117861), 1–14. https://doi.org/10.1016/j.foreco.2019.117861

National Wildfire Coordinating Group. (2021). *Fire Weather Index (FWI) System*.

Plucinski, M., Gould, J., McCarthy, G., & Hollis, J. (2007). The Effectiveness and Efficiency of Aerial Fire Fighting in Australia. *Bushfire Co-Operative Research Centre, Australia*, *Part 1*.

Seneviratne, S. I., Abid, M. A., Pinto, I., & Sylla, M. B. (2021). *IPCC AR6 WGI Chapter11: Weather and Climate Extreme Events in a Changing Climate*. IPCC.

Shahparvari, S., Bodaghi, B., Roozbeh, I., Mohammadi, M., Soleimani, H., & Chhetri, P. (2021). A cooperative (or coordinated) multi-agency response to enhance the effectiveness of aerial bushfire suppression operations. *International Journal of Disaster Risk Reduction*, *61*(102352), 1–15. https://doi.org/10.1016/j.ijdrr.2021.102352

Turco, M., Levin, N., Tessler, N., & Saaroni, H. (2017). Recent changes and relations among drought, vegetation and wildfires in the Eastern Mediterranean: The case of Israel. *Global and Planetary Change*, *151*, 28–35. https://doi.org/10.1016/j.gloplacha.2016.09.002

Zeferino, J. A. (2020). Optimizing the location of aerial resources to combat wildfires: a case study of Portugal. *Natural Hazards*, *100*(3), 1195–1213. https://doi.org/10.1007/s11069-020-03856-6

Zscheischler, J., & Seneviratne, S. I. (2017). Dependence of drivers affects risks associated with compound events. *Science Advances*, *3*(6), 1–11. https://doi.org/10.1126/sciadv.1700263

Towards a Repository for Datasets of Mass Casualty Incidents

Florentina Hager*
University of Twente
f.hager@utwente.nl

Melanie Reuter-Oppermann
University of Twente
m.n.reuter-oppermann@utwente.nl

ABSTRACT

Many mathematical models and algorithms addressing logistics around mass casualty incidents (MCI) have been proposed in the literature. However, hardly any data sets exist that can be used to analyse, validate, and evaluate approaches and solutions. As MCI are comparably rare, real-world data is scarce and often extremely sensitive, making it difficult for researchers to get access and build test datasets. In this paper, we therefore present the first steps towards a repository for the exchange of datasets to facilitate validation and comparison. We first provide an overview of potential stakeholders such as working groups, review disaster-related databases and existing repositories for other fields followed by a preliminary taxonomy to identify and structure the relevant entities and their characteristics within models for disaster management.

Keywords

Dataset, taxonomy, disaster management, operational research, decision support

INTRODUCTION

In 2022, 185 million people were injured, homeless or otherwise affected by disasters (Centre for Research on the Epidemiology of Disasters (CRED) et al. 2022). Providing help to such a high number of people combined with the short or complete lack of time to prepare for a disaster poses logistical challenges. With an increasing number of disasters over the past years, the field of disaster management has received more attention in the literature (Tippong et al. 2022), resulting in a steadily increasing number of published articles that propose models to solve (a subset of) the problems arising in disaster management.

An ongoing problem in developing, validating, and comparing these models is the lack of comprehensive publicly available datasets that include all the necessary information. One reason for the lack of those datasets is the unpredictability of disasters. In contrast to daily operations for which the data collection can be planned in advance, disasters usually occur at no or only short notice and lead to exceptional situations. As a consequence, collecting comprehensive data of good quality is difficult (Shi et al. 2023). Despite these difficulties, several statistics of disasters are publicly available, but these are often not detailed enough to test and validate models. Consequentially, researchers have to make assumptions and educated guesses for a varying range of parameters based on their knowledge of the on-site conditions. From an academic perspective, this approach leads to limited comparability of the models, unclear practical value, and the risk of "over-fitting" the model without the intention to do so.

To overcome this problem, one solution would be to collect, structure, and share existing datasets and add missing information based on expert knowledge. One approach to then provide access to these datasets could be a repository. Such a repository would have multiple advantages. First, it would facilitate access to existing datasets and reduce researchers' workload when testing their model. Second, such a collection of datasets would also facilitate the comparison of actual or assumed values for different parameters across datasets. On the one hand, such a comparison could support the analysis of differences between data sets (e.g. due to different local conditions or disaster types), and, based on this, the general validity of a proposed model. On the other hand, an overview of realistic values for parameters can also support researchers to create, derive, and verify new datasets.

To create a repository that shares the datasets in a structured way, the following steps are necessary (adapted from Leeftink and Hans 2018).

*corresponding author

- **Step 1** Search for existing databases and interest groups.
- **Step 2** Develop a taxonomy to categorise existing models and their datasets.
- **Step 3** Validate the taxonomy with researchers and experts from practice and link the existing datasets and the parameters used therein to the developed taxonomy.
- **Step 4** Derive standardised validation datasets based on existing datasets and discussions with experts.

In this paper, we present the first two steps, while the remaining steps will be part of future works. As the field of disaster management covers a broad spectrum of problems, we will first focus on the subarea of casualty management, which includes models related to managing the needs of casualties, such as treatment, transport, shelter, and relief supplies (Farahani et al. 2020). The remainder of the paper is structured as follows: The next section describes the current work performed for Step 1, followed by a description of Step 2 and a preliminary framework. We conclude the paper with an outlook of future steps and a summary.

STEP 1: WORKING GROUPS, DATABASES, AND REPOSITORIES

To assess whether repositories or validation sets exist, we conducted a literature search and identified and screened potential hosts of these repositories such as working groups, databases and other repositories as examples.

Working Groups in Humanitarian Logistics and Repositories

Worldwide, many different working groups in the area of humanitarian logistics exist. While we focus on Operational Research (OR) / Information system (IS) groups, it is worth noting that input for datasets can also come from medical or practice groups, e.g. the working group on disaster medicine of the German Society for Interdisciplinary Emergency and Acute Medicine (DGINA) (AG Katastrophnemedizin DGINA 2024).

Within the IS community, the Information Systems for Crisis Response and Management (ISCRAM) group is probably the most fitting for our endeavour. As stated on their website (ISCRAM 2024), "the ISCRAM Association's primary mission is to foster a community dedicated to promoting research and development, exchange of knowledge and deployment of information systems for crisis management, including the social, technical and practical aspects of all information and communication systems used in all phases of management of emergencies, disasters and crises." While OR/Management Science models represent a subset within this field, the scope of this working group is broader with respect to the applied methodologies. With yearly conferences and additional events, ISCRAM is one of the main conference outlets for crisis-related research. We also searched within the Association for Information Systems (AIS) and identified the Special Interest Group on IT in Healthcare, AIS SIGHealth, as potentially relevant as well (AIS SIGHealth 2024). In contrast to ISCRAM, the working group does not organise an individual conference, but offers healthcare tracks at the main IS conferences, e.g. ECIS. It seems less suited to host the repository but can still offer valuable access to existing data sets and the chance to advertise a repository, once it exists.

On the OR side, two working groups within the Association of European Operational Research Societies, EURO, are most promising: 1) HOpe, the EURO working group on Humanitarian Operations (EURO working group HOpe 2024), and 2) ORAHS, the EURO Working Group of Operational Research Applied to Health Services (EURO working group ORAHS 2024). The purpose of HOpe "is to create a platform that facilitates communication among the interested parties and forms an active group" (EURO HOpe mini conference 2024) while "Humanitarian Operations Research studies decision-making problems related to providing logistical assistance for humanitarian purposes, and in response to humanitarian crises" (EURO HOpe mini conference 2024). The College of Humanitarian Operations and Crisis Management (HO & CM) of POMS "seeks to create a focus group of researchers and practitioners to study how to manage operational systems under stress or severe resource constraints. A prime objective is to foster a creative environment among academics and practitioners to recognise how operations management plays a crucial role in dealing with the alleviation of human suffering under conditions triggered by cataclysmic acts of man and nature." Within the Institute for Operations Research and the Management Sciences (INFORMS), we also identified two potentially fitting working groups, namely the Health Applications Society (INFORMS Health Applications Society 2024) and the Military and Security Society (INFORMS Military and Security Society 2024), although it is not clear from their website whether they also focus on the humanitarian aspects. Moreover, we identified additional associations, mainly country-specific, that also include humanitarian logistics, e.g. during their yearly conferences. Examples include the Decision Science Institute (DSI), the German OR Society and the UK OR Society.

Based on our investigations we conclude that to the best of our knowledge, a repository for research does not exist on the websites of any of those working groups yet.

Existing Databases

Besides repositories provided by working groups, other sources for input data are existing databases. Within the field of disaster management, various databases exist. A widely cited database with global coverage is the Emergency Event Database (EM-DAT) (Jones et al. 2023) hosted by the Centre for Research on the Epidemiology of Disasters (EM-DAT et al. 2024). The database includes natural and technological disasters and provides information about a disaster's type, region, humanitarian, and economic impact. Data on disasters with a terrorist background is collected by the Global Terrorism Database (LaFree and Dugan 2007). For each recorded incident, the database provides information about the location, the attacker, the method, and the number of casualties. Datasets from various sources are made available on the Humanitarian Data Exchange (HDX, OCHA 2024) platform hosted by the United Nations Office for the Coordination of Humanitarian Affairs (OCHA). Unlike the other two databases, this platform does not only focus on disasters, and datasets are provided by various organisations, resulting in different data structures. Besides these databases, several databases on a national level exist, such as the Canadian Disaster Database (CDD 2024), for example. Although these databases can be useful for getting insights about the size of a disaster, they usually lack the detailed information required for validating models for disaster management, such as locations of included hospitals and shelters, their capacities, or the available fleet of vehicles.

Repositories and Benchmark Sets Outside of Humanitarian Logistics

While we could not identify any existing repositories for humanitarian logistics, there exist repositories or benchmark sets for other problems within the field of OR which can provide insights into the purposeful design of a repository and/or taxonomy. The VRP-REP (Mendoza et al. 2014) is a repository for instances of the Vehicle Routing Problem (VRP). It stores problem variants, references, datasets, and the best-known solution for each combination of problem variants and datasets. Instances for shift scheduling problems can be shared on www.schedulingbenchmarks.org (*Schedulingbenchmarks* 2024). The repository differentiates between instances for nurse rostering, multi-activity multi-day and other employee scheduling problems, and similar to VRP-REP links the different instances, references and best-known solutions. Instances for planning and scheduling of projects are collected and shared by Ringer and Fox (2024). The aim is to provide realistic instances to enhance the development of models with significant practical value. For problems related to operating room scheduling, Leeftink and Hans (2018) developed a benchmark generator to facilitate the access to realistic datasets. Moreover, for each of the different instances as identified by Leeftink and Hans (2018), 200 samples are provided to enhance comparison between different models and, if already found, provide the optimal solution.

These repositories can serve as a guideline for the development of a repository of humanitarian logistics. With the large variety of problems addressed within the literature on disaster management, the repository should provide an overview of existing variants, provide realistic values for the different characteristics, and allow scholars to create a dataset based on their needs.

STEP 2: TAXONOMY

After having identified relevant working groups, existing databases and repositories, the following step is to develop a taxonomy for classification. First, we review the literature for existing frameworks that could be applied to our purpose, before we present our developed taxonomy to describe the relationship between the attributes of different entities.

Previous Taxonomies

In the past years, several taxonomies have been presented within the field of disaster management. On a broader level Kuipers and Welsh (2017) provide a taxonomy to classify the literature. Based on frequent terms and keywords used they derive 19 umbrella terms, thereof ten for crisis types and nine for crisis themes. Wong et al. (2017) present a taxonomy that provides an overview of evaluation typologies. Therefore, they define three layers, the timeline of the disaster, the status, and the overarching theme of resilience, to which the different typologies are related. Rauner et al. (2018) propose a taxonomy that links tasks in disasters to responders. They identify eleven interventions in various phases of the response and three broad categories of responders, further refined for each category. Tatham and Spens (2011) develop a taxonomy to store explicit knowledge in humanitarian logistics. Combining two existing frameworks, the "Supply Chain Operations Reference" and the "Lines of Development", the authors adopt the five primary processes of the supply chain and eight force elements to create a taxonomy. Besides these taxonomies, several other taxonomies have been published. However, these focus on very specific aspects of disaster management, such as simulation tools (Barthe-Delanoë et al. 2015), volunteers (Betke et al. 2023), or IT support (Pottebaum et al. 2014), among others. Taxonomies have also been proposed in related areas like emergency

logistics (Reuter-Oppermann and Bernath 2016; Reuter-Oppermann and Rachuba 2017, for example), which might also provide some relevant insights.

Despite the variety of taxonomies within the literature, we could not find any taxonomy addressing humanitarian logistics on a sufficiently detailed level so that it could serve as a basis for the classification of datasets. Therefore, in the following, we present a preliminary taxonomy for the purpose of classifying datasets.

Preliminary Taxonomy

To develop the taxonomy, we follow the approach suggested by Nickerson et al. (2013) and Kundisch et al. (2022). The purpose of this taxonomy is to identify and structure the relevant entities and their characteristics within models for disaster management. This leads to the meta-characteristic of our taxonomy, namely, to enhance the classification of datasets and the therein included parameters. We set objective and subjective ending conditions and defined the objective ending conditions to be met once all papers of our sample have been examined, and the entities and characteristics have not been changed (i.e. no elements were added, deleted, or updated). For the subjective ending condition, we evaluated whether the proposed taxonomy is concise, robust, comprehensive, extendable, and explanatory. Eventually, as an evaluation goal, the taxonomy should support users to classify and analyse models and datasets within the field of disaster management. We chose an empirical-to-conceptual approach and analysed a selection of papers within the field of casualty management in disasters as presented in Farahani et al. (2020). We opted to use this literature review as a starting point, as it covers the relevant papers within the field until 2019. We acknowledge that by doing so we exclude more recently published papers. However, we believe that the selection of the papers is broad enough to develop a general taxonomy that should also be applicable to papers published after 2019, and if necessary, be extendable to incorporate new characteristics as a next step. From the selected papers, we collected the parameters and sets used in the mathematical models. For the first version of the taxonomy, we derived the entities based on the most common sets in the models and their characteristics based on the indexed parameters. For each entity, we then assessed whether *i)* the entity has unique characteristics or could be merged with another entity, and *ii)* the entity is common enough to be considered a general entity in the taxonomy. We then re-evaluated the attribution of the characteristics, in particular those characteristics that originate from parameters with multiple indices. For each characteristic, we assessed *i)* the dependency, both on entities and on other characteristics, and, similarly to entities, *ii)* their generality. We then used this first version for a (preliminary) validation and assessed the taxonomy's applicability. Based on the results, we modified the taxonomy and repeated the process until the ending conditions were met.

Eventually, we identified ten entities and clustered them into three subcategories, as depicted in Table 1.

Table 1. Entities

Overarching entities	Immobile entities	Mobile entities
Network	Incident sites	Casualties
Budget	Medical centres	Vehicles
	Distribution centres	Commodities
	Shelters	Staff

The characteristics of each of these entities are explained in more detail in the following subsections and are summarised in Table 2.

Network

The network, or graph, formalises the spatial relationship between entities. The locations of immobile entities can be represented as nodes, and the transportation network connecting these locations as arcs (Figure 1). Each node and arc can have several attributes. The attributes of nodes depend on the type of the entity and will be further explained in the respective section. Arcs can be characterised by two main attributes: the distance, i.e. the length of an arc, and the status. The status refers to the condition of an arc. Especially after natural disasters, arcs can be damaged or blocked, leading to a reduced capacity or (partial) unavailability for routing.

Table 2. Overview of characteristics

	Network	Budget	Incident sites	MC	DC	Shelters	Casualties	Vehicles	Commodities	Staff
Distance	✓									
Availability of arcs	✓									
Capacity of arcs	✓									
Free budget		✓								
Designated budget - MC		✓								
Designated budget - DC		✓								
Designated budget - Shelters		✓								
Designated budget - Commodities		✓								
Designated budget - Casualties		✓								
(Initial) Location			✓	✓	✓	✓	✓	✓		✓
Onset time			✓							
Damage			✓							
Required staff			✓		✓		✓			
Required commodities			✓				✓			
Level / type				✓	✓	✓	✓	✓	✓	✓
Capacity				✓	✓	✓		✓		
Variable capacity				✓	✓					
Fixed capacity				✓	✓					
Storage capacity					✓	✓				
Distribution capacity					✓	✓				
Capacity for affected people						✓				
Stock					✓	✓				
Establishment cost				✓	✓	✓				
Extension cost				✓	✓					
Fixed required staff					✓					
Variable required staff					✓					
Number							✓	✓		✓
Priority							✓		✓	
Survival probability							✓			
Maximum waiting time							✓			
Eligible destinations							✓			
Eligible vehicles							✓			
Treatment times				✓			✓			
Transportation cost							✓	✓		
Travel speed								✓		
Eligible arcs								✓		
Facilities for (un)loading								✓		
Time for (un)loading								✓		
Supply / Availability									✓	
Supplier									✓	
Dimension									✓	
Unit cost									✓	
Holding cost									✓	
Shortage penalty									✓	
Qualification										✓
Allowed tasks										✓
Skill level										✓
Wages										✓
Availability										✓
Shifts										✓

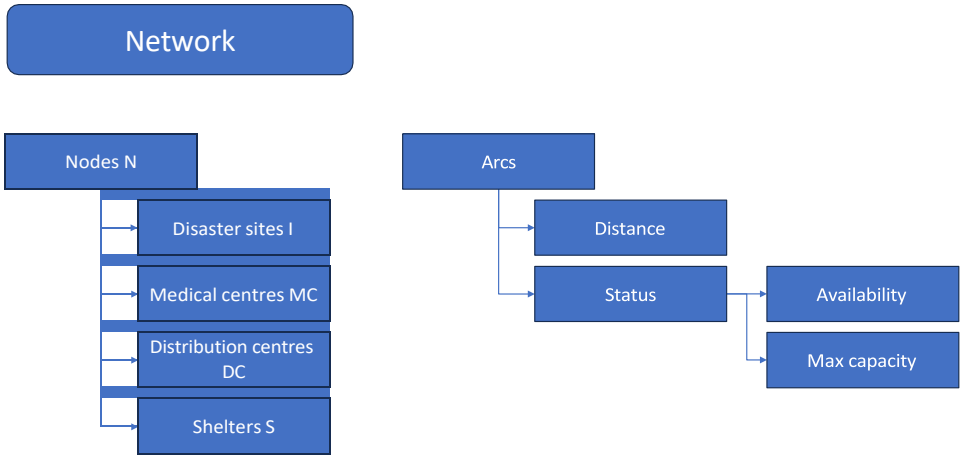


Figure 1. Network

Budget

The budget includes the financial funds to manage the disaster. It can be divided into designated and free budgets (Figure 2). Free budget can be used for any disaster response relief measure, while designated budget is reserved for specific measures, such as constructing shelters, medical centres or distribution centres or the purchase of commodities. Although usually not referred to as budget, a frequently used alternative is to define the number of potential distribution centres/medical centres/shelters to be located or commodities to be bought. This slightly different definition of budget by dividing the monetary budget by the costs for construction or purchase may facilitate the interpretation of these values.

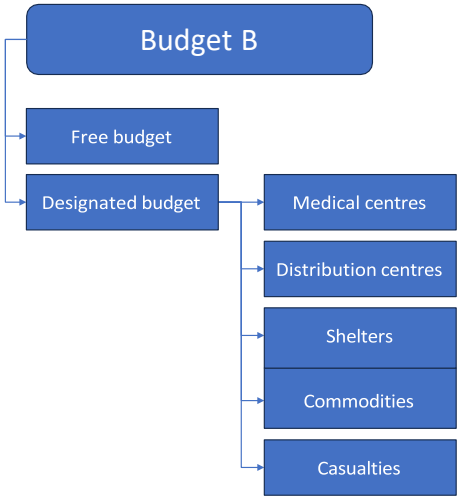


Figure 2. Budget

Incident Sites

The most important nodes in the network relevant to disaster management, although often not explicitly mentioned, are the incident sites (Figure 3). These incident sites are primarily defined by their geographical location, usually defined by a single pair of coordinates for a central point such as a point of registration. Each location can differ in terms of the disaster's onset and the resulting damage. Combined, these characteristics determine, among others, the number of casualties and affected people or the status of the close transportation network at a certain moment. To be able to respond to a disaster, each incident site may require different resources such as commodities and staff. While some of these resources depend on the number and type of casualties and are therefore linked to the entity of casualties, other resources depend mainly on the location of the disaster. Examples of such location-dependent

resources are staff required for general tasks as search and rescue or managing the disaster site and equipment or commodities not linked to specific entities.

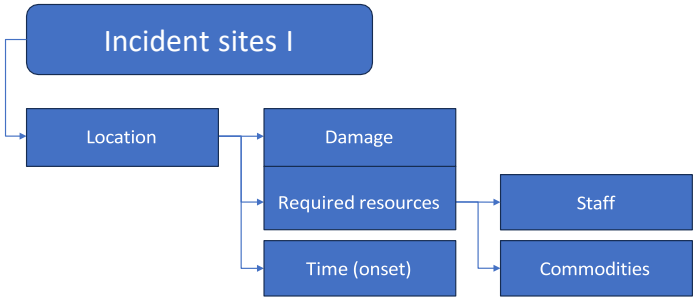


Figure 3. Incident sites

Medical Centres

A medical centre is an entity that provides treatment to casualties (Figure 4). The location of this entity can either be predefined or to be selected from a subset of potential locations. Medical centres can vary in their level, ranging from top-tier hospitals to temporary care units. A level is characterised by the capacity to treat casualties, which can be a fixed capacity or extendable as needed. Moreover, the level of a medical centre can impact the treatment times due to the equipment and the general experience of their staff. These treatment times may further depend on additional factors such as the casualty type or the skill of the treating staff. However, these factors are not explicitly mentioned in the overview to keep the taxonomy as general as possible. Beyond the level's impact on treatment-related attributes, the differences in size and equipment can also be reflected in the establishment and expansion costs.

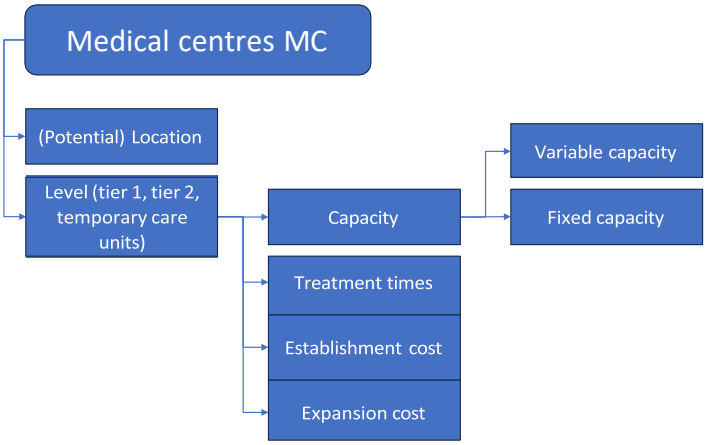


Figure 4. Medical centres

Distribution Centres

Distribution centres (DC) are entities that store and distribute commodities (Figure 5). Their location may be predetermined or selected by the disaster manager from a set of potential locations. DC can also differ in their level, ranging from national DC to temporary local DC, each level impacting various attributes. The most important attributes include the capacity to store and distribute commodities and the current stock. Similar to medical centres, the capacity can be fixed or extendable. While these attributes describing the capacity can be further specified, such as by commodity type, for the sake of generality, these extensions are not explicitly listed in the overview. To operate a DC of a certain level, staff is necessary, which can be further specified by the fixed staff requirement, e.g. for administrative tasks, and the capacity-dependent staff requirement. The differences among DC levels are also reflected in the associated establishment and extension costs.

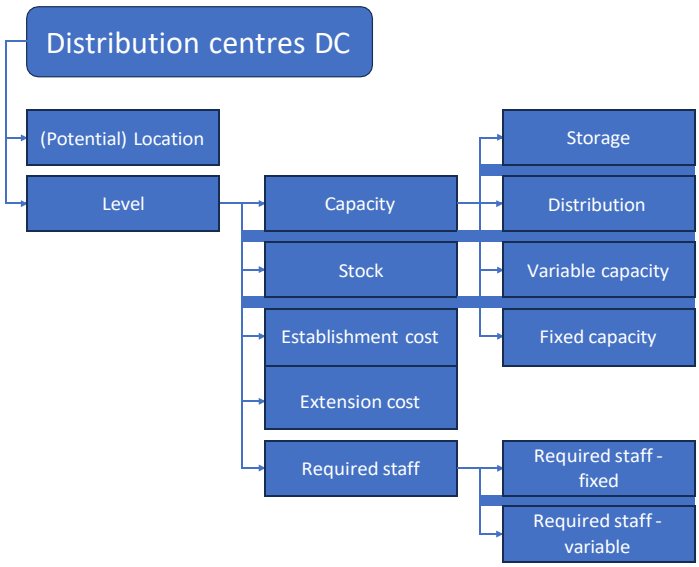


Figure 5. Distribution centres

Shelters

In the literature of disaster management, shelters are entities that can handle commodities and provide shelter to affected people. Hence, they share some characteristics with DC and medical centres (Figure 6). Their location is already predetermined or must be selected by the decision maker. Shelters can be characterised by their type, which may vary in terms of their capacity, stock, and the cost of building a new shelter. Similar to DC, capacity refers to the storage and distribution of commodities. Moreover, as shelters also provide a safe place for those affected, another characteristic of capacity is the number of people that can be accommodated in a shelter.

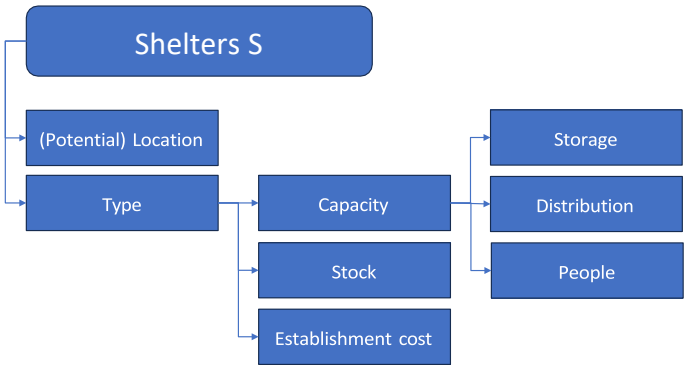


Figure 6. Shelters

Casualties

Casualties are the most important entity in casualty management (Figure 7). This entity usually has three main attributes, namely the type, e.g. the injury types or triage categories, the initial location, which is usually one or more incident sites, and the number of casualties, which is defined either by a fixed number or an arrival rate, e.g. to represent rescue/extrication. The type of casualty has an impact on several other attributes. First, it can determine the priority and the survival probability function or the transition probability from one type to another, the latter of which can be interpreted as a discrete form of the survival probability function. Regardless of the formulation, both can determine the maximum waiting time for a casualty. At the incident site, casualties may need treatment, and the treatment time depends on their type. Different types of casualties may also require different additional resources, resulting in different eligible medical centres or shelters, eligible vehicles, required commodities, such as water, food, or blankets, or required staff per time unit.

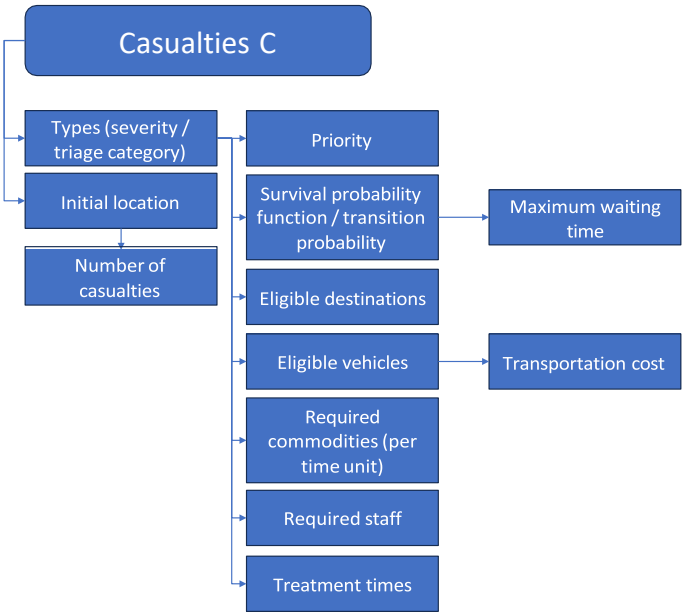


Figure 7. Casualties

Vehicles

Vehicles are entities for transporting casualties, commodities, and staff, and can be characterised by the location, the number of vehicles available and the type (Figure 8). Vehicle types can differ in their travelling speed and eligible arcs, which together with the distance of those arcs determine the travel time between two locations. In addition to the travel time, loading and unloading the casualties and commodities may require additional time or specific facilities. The quantity of commodities and people that can be transported is determined by the capacity of a vehicle type and can be defined by the eligible load, the weight, or the volume of commodities or casualties. Transport costs vary among vehicle types and may also be influenced by the specific load type. However, for the sake of generality, this detailed specification is not included in the overview.

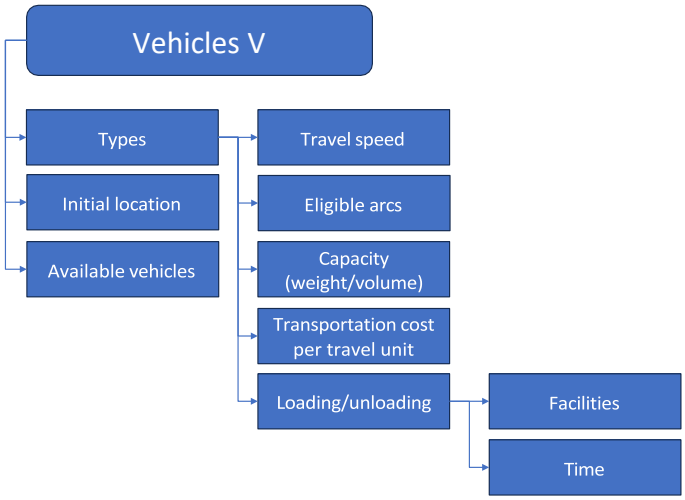


Figure 8. Vehicles

Commodities

To supply affected people and casualties, various types of commodities are required. These types can differ in their priority, dimensions, associated costs, and availability (Figure 9). Costs can include unit costs, holding costs and penalties for shortage (other potential cost types are not included in the taxonomy for reasons of generality). Availability can further be specified on a supplier level.

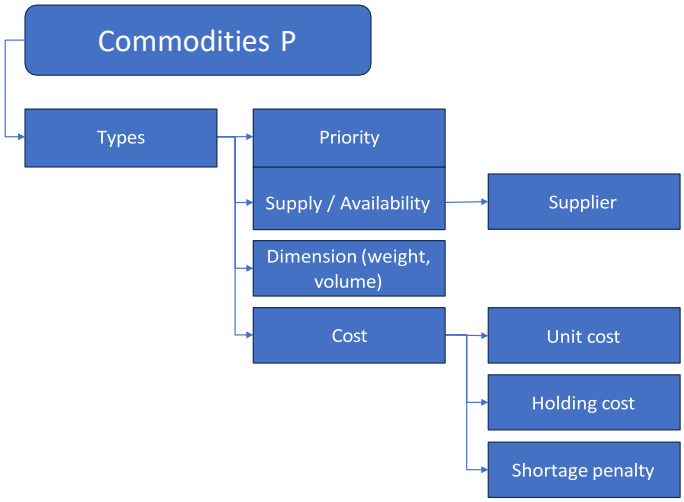


Figure 9. Commodities

Staff

Staff is an entity that treats patients, handles commodities, or fulfils other necessary tasks (Figure 10). It includes paid staff as well as (spontaneous) volunteers. Similar to casualties and vehicles, staff can be characterised by their type, location, and number. The number of staff can either be static or change over time due to availability and resulting shifts. The type of staff can affect the associated wages and can further be specified by the qualification. The qualification may determine the tasks staff members are allowed to perform and can be further refined by the skill level, which in turn may affect the quality and time required to complete a task.

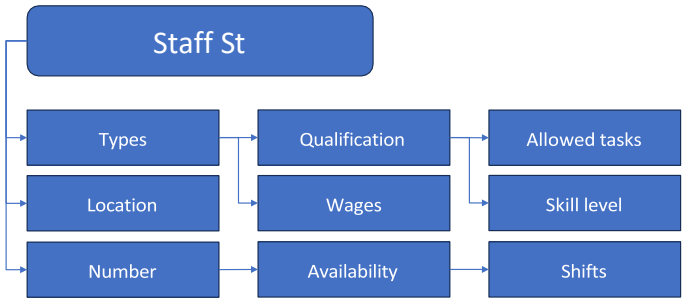


Figure 10. Staff

NEXT STEPS

The next step is to validate the presented taxonomy (Szopinski et al. 2019). For this process, we will invite academics and practitioners with strong expertise in disaster management to participate in discussion rounds, focus groups or interviews. We also aim to discuss the inclusion of entities and characteristics that are not or only implicitly mentioned in the literature and therefore are not (yet) included in the taxonomy. For instance, in practice, general circumstances such as the weather or the region might impact the characteristics of various entities. While they are generally not described in the literature, they could be added to the taxonomy as an additional overarching entity for completeness. Another point for discussion could be the definition of a disaster site. Currently, scholars define them as a set of single coordinates representing one specific point within a disaster area, such as a casualty collection point. However, depending on the disaster, single points may not be the best representation, and instead of defining a discrete set of locations, defining an area might be more representative. Furthermore, to facilitate the use of the repository, clear minimum requirements for the metadata should be defined. For instance, although certain disaster characteristics do not directly affect the model parameters (e.g. the specific cause of building damage) or are already implicitly included in the entity characteristics (e.g. the size of the affected area represented by the number of nodes and distances) and are therefore not included in the taxonomy, adding such information could be valuable for researchers to validate and compare their models.

After the evaluation with experts, we aim to map the published data (sets) and parameters to the taxonomy from both, articles used for taxonomy building, i.e. the articles selected by Farahani et al. (2020), and more recent articles. These two approaches should allow us to obtain an overview of available instances, which can be based on either real or hypothetical data. Based on this, we aim to create benchmark datasets in collaboration and discussion with experts and publish them in a repository to facilitate the comparison of existing and validation of future models in the field of disaster management.

CONCLUSION

In this paper, we proposed a road map for the creation of a repository for the sharing of humanitarian logistics datasets. We provided an overview of currently existing working groups and benchmark sets and presented a preliminary taxonomy for the field of casualty management. These two steps serve as a basis for the future set-up of a repository. Apart from the steps outlined in the previous subsection, future research could focus on extending the taxonomy to other areas of disaster management that are not yet covered to expand the usability of such a repository.

REFERENCES

AG Katastrophnemedizin DGINA (2024). url: https://www.dgina.de/arbeitsgruppen/ag-katastrophenmedizin (visited on 01/29/2024).

AIS SIGHealth (2024). url: http://www.aissighealth.com/wordpress/ (visited on 01/29/2024).

Barthe-Delanoe¨, A.-M., Truptil, S., and Benaben, F. (2015). "Towards a taxonomy of crisis management simulation tools". In: *Proceedings of the 12th International ISCRAM Conference*. ISCRAM, 7–p.

Betke, H., Bosse, M., Sackmann, S., and Reusch, C. (2023). "Towards a Taxonomy for Classification of Coordination Systems for Spontaneous Volunteers". In: *Proceedings of the 20th International ISCRAM Conference*. ISCRAM, pp. 751–760.

CDD (2024). *Canadian Disaster Database*. url: https://www.publicsafety.gc.ca/cnt/rsrcs/cndn-dsstr-dtbs/index-en.aspx (visited on 01/25/2024).

Centre for Research on the Epidemiology of Disasters (CRED), Universite´ Catholique de Louvain (UCLouvain), and United States Agency for International Development (USAID) (2022). *Disasters in numbers*. https://www.cred.be/sites/default/files/2022_EMDAT_report.pdf (visited on 03/27/2024).

EM-DAT, Guha-Sapir, D., Below, R., and Hoyois, P. (2024). *The CRED/OFDA International Disaster Database*. url: www.emdat.be (visited on 01/25/2024).

EURO HOpe mini conference (2024). url: https://www.bath.ac.uk/events/euro-hope-mini-conference-2023/ (visited on 01/29/2024).

EURO working group HOpe (2024). url: https://www.euro-online.org/web/ewg/39/euro-working-group-on-humanitarian-operations-hope (visited on 01/29/2024).

EURO working group ORAHS (2024). url: https://orahs.di.unito.it/ (visited on 01/29/2024).

Farahani, R. Z., Lotfi, M., Baghaian, A., Ruiz, R., and Rezapour, S. (2020). "Mass casualty management in disaster scene: A systematic review of OR&MS research in humanitarian operations". In: *European Journal of Operational Research* 287.3, pp. 787–819.

INFORMS Health Applications Society (2024). url: https://connect.informs.org/healthapplications/home (visited on 01/29/2024).

INFORMS Military and Security Society (2024). url: https://connect.informs.org/militaryandsecurity/home (visited on 01/29/2024).

ISCRAM (2024). url: https://iscram.org/ (visited on 01/29/2024).

Jones, R. L., Kharb, A., and Tubeuf, S. (2023). "The untold story of missing data in disaster research: a systematic review of the empirical literature utilising the Emergency Events Database (EM-DAT)." In: *Environmental Research Letters* 18.10, p. 103006.

Kuipers, S. and Welsh, N. H. (2017). "Taxonomy of the crisis and disaster literature: Themes and types in 34 years of research". In: *Risk, Hazards & Crisis in Public Policy* 8.4, pp. 272–283.

Kundisch, D., Muntermann, J., Oberla¨nder, A. M., Rau, D., Ro¨glinger, M., Schoormann, T., and Szopinski, D. (2022). "An update for taxonomy designers: methodological guidance from information systems research". In: *Business & Information Systems Engineering* 64, pp. 421–439.

LaFree, G. and Dugan, L. (2007). "Introducing the global terrorism database". In: *Terrorism and political violence* 19, pp. 181–204.

Leeftink, G. and Hans, E. W. (2018). "Case mix classification and a benchmark set for surgery scheduling". In: *Journal of scheduling* 21, pp. 17–33.

Mendoza, J., Gue´ret, C., Hoskins, M., Lobit, H., Pillac, V., Vidal, T., and Vigo, D. (2014). "VRP-REP: the vehicle routing community repository". In: *Third meeting of the EURO Working Group on Vehicle Routing and Logistics Optimization (VeRoLog)*.

Nickerson, R. C., Varshney, U., and Muntermann, J. (2013). "A method for taxonomy development and its application in information systems". In: *European Journal of Information Systems* 22, pp. 336–359.

OCHA (2024). *Humanitarian Data Exchange*. url: https://data.humdata.org (visited on 01/25/2024).

Pottebaum, J., Marterer, R., and Schneider, S. (2014). "Taxonomy of IT support for training emergency response & management." In: *Proceedings of the 11th International ISCRAM Conference*. ISCRAM, pp. 379–383.

Rauner, M. S., Niessner, H., Odd, S., Pope, A., Neville, K., O'Riordan, S., Sasse, L., and Tomic, K. (2018). "An advanced decision support system for European disaster management: the feature of the skills taxonomy". In: *Central European Journal of Operations Research* 26, pp. 485–530.

Reuter-Oppermann, M. and Bernath, C. (2016). "German data sets for comparing ambulance location models". In: *Proceedings of the Second Karlsruhe Service Summit Research Workshop*, pp. 102–109.

Reuter-Oppermann, M. and Rachuba, S. (2017). "Towards a Testing Framework for Location Problems in EMS Logistics". In: *Proceedings of the KSS Research Workshop: A Selection of Talks and Presentations on Designing the Digital Transformation*.

Ringer, M. and Fox, B. (2024). *Planning & Scheduling Benchmarks*. url: https://www.cs.jhu.edu/~jason/325/hw2/benchmrx/default.html (visited on 01/29/2024).

Schedulingbenchmarks (2024). url: http://www.schedulingbenchmarks.org/ (visited on 01/26/2024).

Shi, Y., Liu, N., and Wan, G. (2023). "Treatment Planning for Victims with Heterogeneous Time Sensitivities in Mass Casualty Incidents". In: *Operations Research*.

Szopinski, D., Schoormann, T., and Kundisch, D. (2019). "Because Your Taxonomy is Worth It: Towards a Framework for Taxonomy Evaluation." In: *Proceedings of the 27th European Conference on Information Systems (ECIS)*.

Tatham, P. and Spens, K. (2011). "Towards a humanitarian logistics knowledge management system". In: *Disaster Prevention and Management: An International Journal* 20.1, pp. 6–26.

Tippong, D., Petrovic, S., and Akbari, V. (2022). "A review of applications of operational research in healthcare coordination in disaster management". In: *European Journal of Operational Research* 301.1, pp. 1–17.

Wong, D. F., Spencer, C., Boyd, L., Burkle, F. M., and Archer, F. (2017). "Disaster metrics: a comprehensive framework for disaster evaluation typologies". In: *Prehospital and disaster medicine* 32.5, pp. 501–514.

Cascading Effects of Critical Infrastructures in a Flood Scenario: A Case Study in the City of Cologne

Moritz Schneider
German Aerospace Center (DLR), Institute for the Protection of Terrestrial Infrastructures, Germany
moritz.schneider@dlr.de

Peter Priesmeier
Cologne University of Applied Sciences, Institute of Rescue Engineering and Civil Protection, Germany
peter.priesmeier@th-koeln.de

Alexander Fekete
Cologne University of Applied Sciences, Institute of Rescue Engineering and Civil Protection, Germany
alexander.fekete@th-koeln.de

Daniel Lichte
German Aerospace Center (DLR), Institute for the Protection of Terrestrial Infrastructures, Germany
daniel.lichte@dlr.de

Frank Fiedrich
Chair for Public Safety and Emergency Management, University of Wuppertal, Germany
fiedrich@uni-wuppertal.de

ABSTRACT

Critical infrastructures, which constitute the backbone of our modern society, are increasingly exposed to natural hazards. Loss of performance or failure of a critical infrastructure can lead to cascading effects that affect even more services and citizens. Floods, as one of the most prominent natural hazards, are prone to affect multiple critical infrastructures at once, making it even more difficult to assess combined effects of these (cascading) disruptions. Hospitals are especially vulnerable in a flood scenario, as they are reliant on multiple other infrastructures, such as power and water supply or the road network. In order to prepare for upcoming events, sophisticated analysis tools are required that are capable of modeling the spatial extent of flood induced disruptions and their impact on critical infrastructure services. In this work, we present a proof of concept that focuses on the impact of multiple disruptions on hospitals. We conducted a case study on an extreme flood scenario in the city of Cologne (Germany). Historically, Cologne has proven vulnerable to river-floods, as thousands of people were affected through floods in in 1993 and 1995. The approach is based on a combination of I) a geographic information system, which makes the extent of disruptions spatially explicit and II) a Bayesian network, which is used to assess the impact of one or multiple disruptions on a single hospital. We present a work-in-progress approach in this work. The results generated using this approach enable a first comparative overview of the expected level of services of the examined hospitals.

Keywords

Critical Infrastructure, Cascading Effects, Flood Risk Management, Bayesian Network, GIS.

INTRODUCTION

Growing interdependencies of critical infrastructures (CI) in combination with increasing exposure to natural hazards contribute to society's vulnerability to CI breakdowns (Nick et al., 2023). In addition, the functionality of one CI is most often reliant on the functionality of at least one other CI (Rinaldi et al., 2001), which poses a challenge for disaster resilience. The resulting highly interconnected networks can lead to severe cascading effects in case of a single disruptive event, such as a flood (Fekete, 2019). These interdependencies can cause impacts on a local CI that can have far-reaching effects on other CI. For this reason, it is important to have an in-depth understanding of disruptions and the resulting cascading effects to better prepare for upcoming challenges. This requires sophisticated CI analysis methods that look at exposure to natural hazards and also investigate CI interdependencies enabling a prediction of cascading effects and their impacts.

In this work, we study how the impact of multiple (cascading) flood-induced disruptions on different service levels of a CI can be analyzed for flood risk management. In terms of CI, we focus on hospitals as they are crucial for disaster response and especially vulnerable to floods and failure of other CI services (Melnychuk et al., 2022). We thus require a model that (1) makes the distribution of the considered disruptions spatially explicit, and (2) enables an estimation of the effect of one or multiple disruptions on predefined service levels of a hospital. A case study of an extreme flood scenario in the city of Cologne (Germany) is conducted as a proof of concept. The final goal of the approach is to create a map that provides a comparative overview of the expected level of services between the examined hospitals.

Here, we present an approach to address these goals. The main contribution of this work is an approach for assessing the impact of multiple flood-induced cascading effects (i.e. disruptions) on hospitals. The novelty of this approach is the consideration of individual as well as combined disruptions on hospitals in a flood scenario taking into account uncertainties. We argue that it is important to consider the uncertainties both of the disruptions as well as in their impact on a hospital. As a result, the approach generates a map which can be used to compare affected hospitals to one another serving as an information tool for emergency responders in flood risk management. Especially in emergency response, it is important to consider cascading sequences triggered by natural events (Ricci et al., 2024).

The approach is based on a combination of a geographic information system (GIS) and a Bayesian network (BN) - the GIS informs the BN (as e.g. outlined in Johnson et al., 2012). The GIS is used to model the spatial extent of the considered disruptions. Geographic information systems are used extensively in emergency and crisis management, including impact analysis of natural hazards (e.g. see Tzavella et al., 2018 or Geiß et al., 2022). The Bayesian network is used to model the impact of the disruption(s) on a single hospital under consideration of uncertainties. Bayesian networks, which are probabilistic graphical models (Pearl, 1988), are often used in the literature to model the effects of natural hazards on CI (e.g. see Kameshwar et al., 2019 or Ram´ırez-Agudelo et al., 2021).

The definition of a disruption in this work follows the proposal by Mentges et al., 2023: A disruption (e.g. a damaged road) is caused by a disruptive event (e.g. a flood) and lasts as long as the performance of the system is decreased as a consequence of the event. The disruption ends when the system has recovered and can thus last longer as the disruptive event, i.e. the flood is gone but the road is still damaged. The concrete disruptions considered in this work are: the flood itself - in terms of a rising water levels, the flooding of hospital access roads (a indirect, i.e. cascading, effect), and potential blackouts caused by the flood (a second cascading effect). We selected these disruptions as the starting point for our work-in-progress approach because they are significant in their impact and are likely to occur in a flood scenario.

The remainder of this paper is structured as followed: First, our work-in-progress approach is described, including the case study of a extreme flood scenario in Cologne, Germany. Subsequently, preliminary results are presented, followed by a discussion and conclusion.

APPROACH

The approach can be divided into three steps: (1) the considered disruptions (flooded area, access road availability, and blackout areas) are modeled in a GIS environment, (2) a Bayesian network model is developed that is used to infer the effect of a single or multiple disruptions on the service levels of a hospital, and (3) the GIS and BN are combined (GIS informs the BN) to estimate the expected level of service of each individual hospital in an automated manner. In the following, the case study is presented, followed by the GIS model and its method to estimate the spatial extent of the considered disruption. Subsequently, the BN is developed and the combination of the GIS and the BN outlined.

Case Study Description

Our approach is developed and tested in the case study area of Cologne, Germany. Cologne is Germany's fourth largest city with a population of about one million people. The city is divided by the Rhine river, which is the second largest river in Germany. There have been several devastating Rhine floods in the past, for example in 1993 and 1995 (Herget and Meurs, 2010), the latter affecting 33,000 people and causing damage of 35 million euros in Cologne alone (Fink et al., 1996). A five-hundred-year flood (extreme flood) hazard map is used as exposure data in this case study (see Figure 1). Three disruptions that can be expected in a flood scenario and affect a hospital are considered in this work-in-progress approach: (1) the flood itself in terms of the rising water level, (2) a potential flooding of access roads to a hospital, and (3) potential blackouts caused by flooded transformer stations or safety shutdowns.

Open Street Map (OSM) data is used to determine access roads of hospitals in Cologne, the location of transformer stations used to determine blackout areas, and the location of hospitals. This data represents example data which was not reviewed for actuality and accuracy. The data of the extreme flood scenario called HQ 500 is provided by official sources (geoportal.nrw, 2024; geoportal of the state of NRW: www.geoportal.nrw/). In this scenario, the effect of failed or overtopped dikes, mobile flood defenses, and intrusion of ground water into old river arms is included (Fekete, 2020).

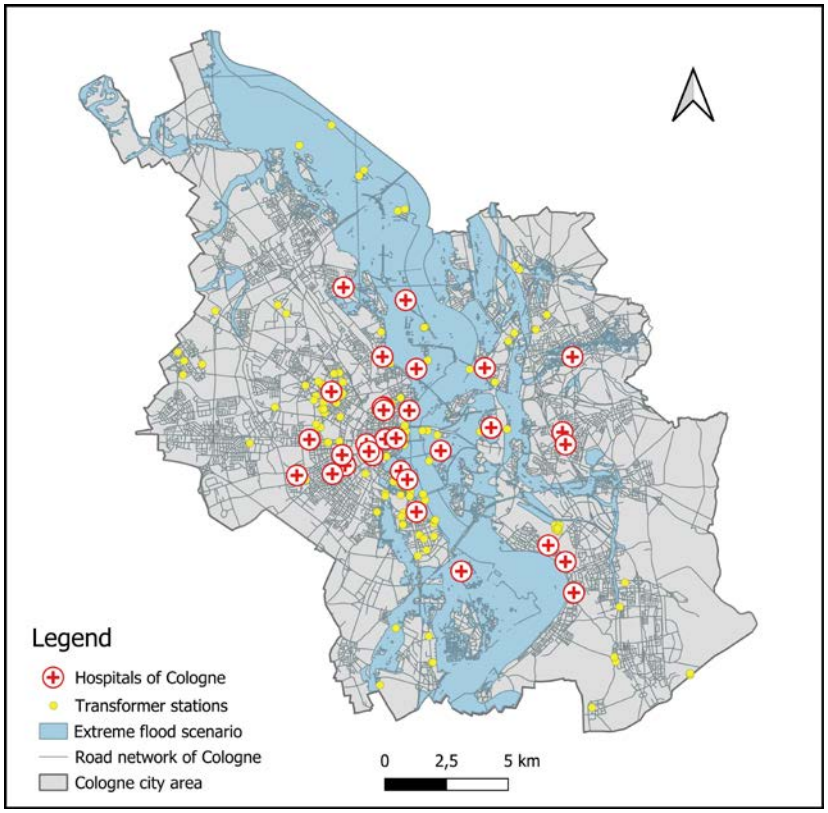


Figure 1. Map of Cologne including the extent of the assumed flood scenario, as well as the road network and locations of transformer stations and hospitals.

GIS Model

The spatial extent of the considered disruptions (flood, blackout, and road availability) is modeled in QGIS, which is an open source geographic information system. The approaches used to assess the spatial extent of these disruptions are work-in-progress in terms of their level of detail, but sufficient for a proof of concept of the approach. In the following, a brief description of the three individual approaches - one for each disruption - is presented.

The *flood* disruption is modeled by examining whether a hospital is located within the flood propagation layer in the GIS. It is analyzed, whether the flood propagation layer shows an overlap with a hospital building (compare Figure 1). In this work-in-progress approach, we conduct a binary assessment that omits the water level.

In order to determine the *access road availability* for each hospital, a ring buffer with a diameter of 100 meters is established around each center point of a hospital. Each access road within this buffer that does not show an overlap with the flood propagation layer is identified (see Figure 2). For each buffer, i.e. each hospital, it is determined if no access road, a single access road, or multiple access roads are available.



Figure 2. Example of a 100 meter buffer around a hospital showing multiple flooded access roads and one access road which can still be used.

The approach to determine blackout areas is based on a method presented by Pala et al., 2014. The method uses on Thiessen polygons (also called Voronoi diagram) without weights. In order to determine the Thiessen polgyons, an algorithm is used that grows cells (that constitute the polygons) in a raster format starting from each source point (Pala et al., 2014), which are the location points of the transformer stations in Cologne. In a next step, it is analyzed which transformer stations are within the flooded area and thus would potentially be damaged or shut down for safety reasons. The corresponding cells to these source points are considered as potential blackout areas (see Figure 3). Four different areas of different blackout probability are considered in the approach: (1) the area that shows an overlap with the flood propagation layer and the affected Voronoi cells is considered as area with highest the blackout probability (assuming a 90% probability), (2) the area that shows an overlap with the flood propagation layer but not with the affected Voronoi cells is considered as the area of medium blackout probability (70% probability), (3) the area that shows an overlap with the affected Voronoi cells, but not with the flood propagation layer is considered as an area of lower blackout probability (50%), and (4) the area that does not show an overlap with both the flood propagation layer and the affected Voronoi cells is considered as the area with zero blackout probability.

Bayesian Network

The Bayesian network (BN) is used in the proposed approach to estimate the impact of one or multiple disruptions on different service levels of each hospital. The BN is implemented in Python based on the pgmpy library (see Ankan and Panda, 2015), which is a versatile library for working with probabilistic graphical models. A BN, introduced by Pearl, 1988, is a directed acyclic graph composed of nodes representing system variables and edges representing conditional dependencies (for a brief introduction of BN see e.g. Puga et al., 2015). Nodes that interact, are connected by edges in the direction of influence. For example, an edge $X \rightarrow Y$ implies that X (the parent node) influences Y (the child node). Additionally, each node has a probability distribution attached to it that defines the chance of finding the node in a given state. Nodes that are not influenced by other nodes (no edge is pointing towards these nodes) are called marginal nodes. Nodes that are influenced by at least one other node (edge(s) pointing towards the node) are called conditional nodes. Conditional nodes have conditional probabilities that are used to infer the state of a child node given information on the states of one or multiple parent nodes.

In our approach, a variable describing the expected *Level of Service* of a hospital is the target child node of the BN. The states of this node constitute potential level of services of a hospital. At this point, it must be considered that the number of states enhances the complexity when defining conditional probabilities for this variable - for each combination of parent and child node states, one probability value must be determined. We selected a simple categorization of hospital service levels that is proposed in Witte et al., 2021: (1) *normal*, which describes the unimpaired operation of the hospital, (2) *emergency*, which describes that the hospital only performs the basic services, and (3) *failure*, which describes that the hospital should be evacuated.

The node *Level of Service* of a hospital is influenced by the considered disruptions. Thus, the three modeled disruptions (flood, road accessibility, and blackout) constitute the parent nodes of node *Level of Service* (see Figure 4). The dependencies between the disruptions are not modeled in the BN, but in the GIS model, thus all disruption

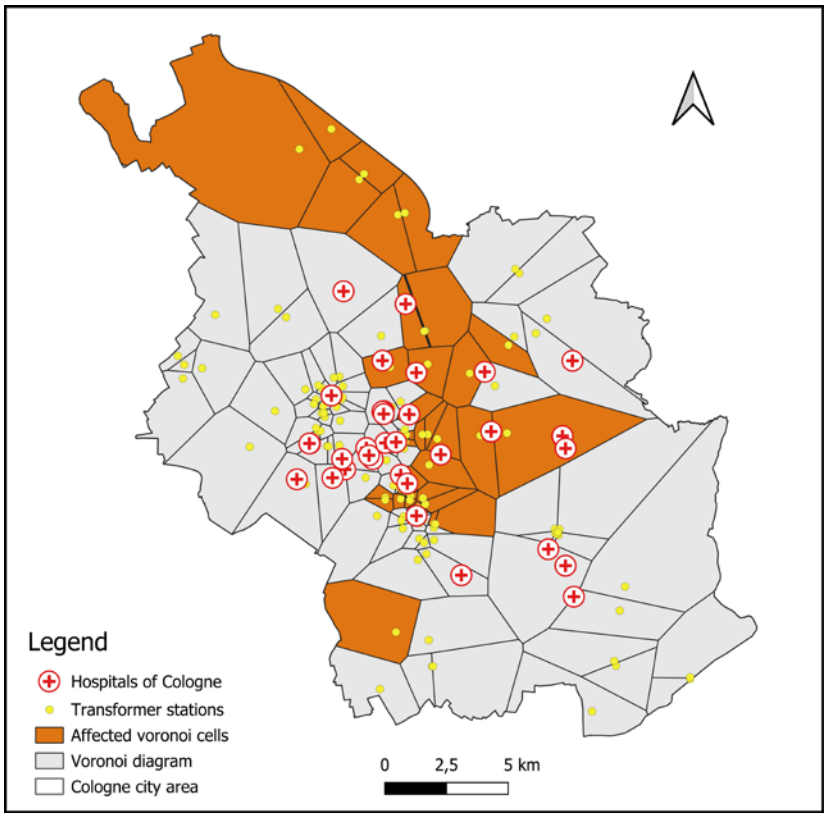


Figure 3. Map of Cologne including the transformer station-based Voronoi diagram including affected Voronoi cells that indicate areas prone to blackouts.

nodes are marginal nodes. The states of the disruption nodes are basic divisions. The nodes *Blackout* and *In Flooded Area* show binary states: *True* and *False*. The node *Road Accessibility* shows three states that describe if no access road is available, a single access road is available, or multiple access roads are available (all nodes and node states are listed in Table 1). The probability values quantifying the nodes of the BN were selected by the authors and reflect the basic dependencies between the variables - the more a hospital is exposed to disruption(s), the more likely it is to be in a state of *failure*. If a hospital is exposed to at least one disruption, the node *Level of Service* cannot be in state *normal*.

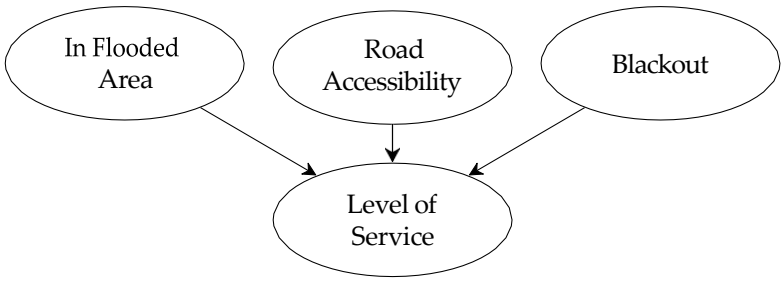


Figure 4. Bayesian network of the approach including the target node *Level of Service* and three parent nodes that describe the considered disruptions.

Table 1. All four considered variables of the Bayesian network shown in Figure 4 including their states.

Variable	**Level of Service**	**Road Accessibility**	**Blackout**	**In Flooded Area**
State 1	Normal	None	True	True
State 2	Emergency	Single Road	False	False
State 3	Failure	Multiple Roads		

In order to calculate the posterior probability of node *Level of Service*, i.e. the probability given evidence on the disruption nodes, the BN (see Figure 4) is duplicated and assigned to each hospital. The individual BNs are informed by the results of the GIS model, i.e. it is analyzed, if a hospital is within the flooded area, if it is in one of the four blackout zones (high, medium, low, and zero probability zone), and how many access roads are available.

Two types of evidence (information about node states) gained from the GIS must be distinguished. The evidence on variable *In Flooded Area* and *road Accessibility* is considered as hard evidence - which describes that we know the exact state of this variable (within the boundaries of the model). For node *Blackout*, we defined four areas with varying blackout probability, thus we do not have evidence with certainty, but evidence of a probability distribution. In order to consider this uncertainty evidence, soft evidence, introduced by Valtorta et al., 2002, is applied. In this way, the posterior probability of node *Blackout* being in state *True* follows the probability assigned to the individual blackout zones.

RESULTS

The results (see Figure 5) are illustrated using the geopandas library in Python that enables geospatial data analysis and visualization (Jordahl et al., 2020). The BN is duplicated and assigned to each of the 35 hospitals and spatially informed by the GIS model. This results in 35 probability distributions of node *Level of Service*. Thus, for each hospital, one probability value for each considered level of service (normal, emergency, and failure) is calculated, which sum up to one. Figure 5 shows the results on three maps. Each map shows the probability of one node state (normal, emergency, failure) for each hospital. Before going into detail on the results, it must be pointed out that these results have limitations due to assumptions within the model environment and cannot be used directly to estimate realistic effects of a flood. Nevertheless, the results allow an assessment of the suitability of the approach in regard to the research objective.

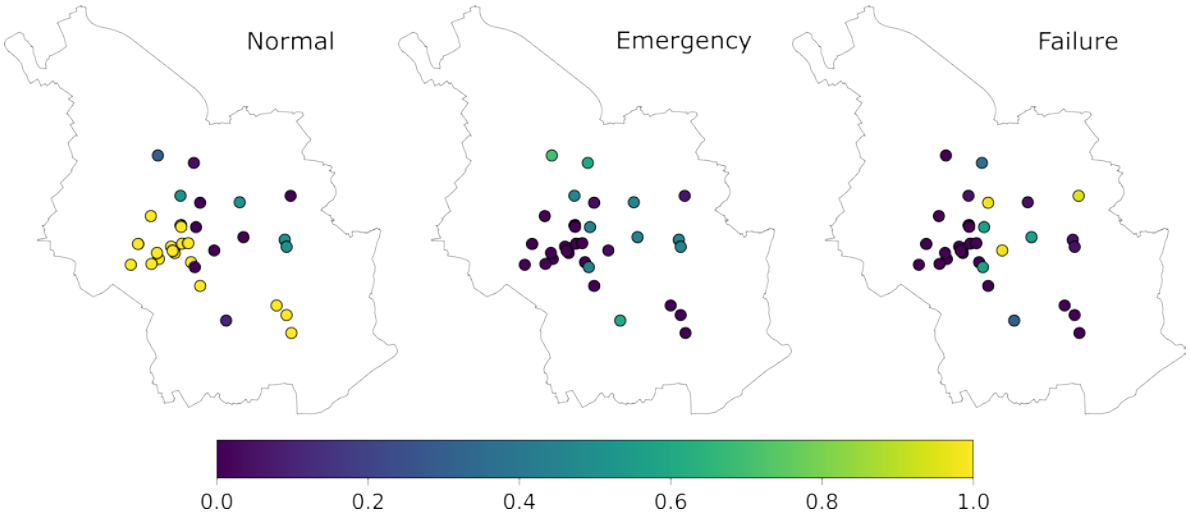


Figure 5. Preliminary results of the case study presented as three maps that illustrate the probability values of the three states of node *Level of Service*. The color of the dots (each representing one hospital) follow the colorbar at the bottom of the figure, thus, each color represents a certain probability value. Looking at one hospital on all three maps, the probability values sum up to one.

The preliminary results show that 20 hospitals are not affected by one of the three considered disruptions, i.e. $P(\mathit{Level\ of\ Service} = \mathit{Normal}) = 1$ (see bright yellow-colored circles in the left image in Figure 5). 17 hospitals are located west of the Rhine river and three hospitals are located in the south east of Cologne. Seven hospitals show a probability of 0% for state *Normal*, i.e. they are affected by at least one disruption. Two of these hospitals show the highest probability of state *Failure* with 97% (see bright yellow-colored circles on the right image in Figure 5). These hospitals are located close to the Rhine river and are within the flooded area as well as within a potential blackout area (leading to the highest probability of a blackout), and no access roads are available. One additional hospital shows a high probability of state *Failure* with 94%. This hospital is located near a tributary of the Rhine and is located within the flooded area and an affected Voronoi cell, but there is one access road still available.

DISCUSSION AND CONCLUSION

Cascading effects are of increasing interest in research, not only in the context of supply chains, but also with regard to the interconnection of natural hazards and related impact chains. However, concrete assessments, showcasing

how these interactions and cascading effects play out are still needed. The presented work-in-progress approach provides a first step to assess cascading effects and resulting multiple disruptions on hospitals in a flood scenario. By introducing a BN, the effect of a single or multiple combined disruptions is modeled including consideration of uncertainties in both the intensity of the disruption and the effect on different service levels of a hospital. Using a GIS model, the spatial extent of the disruptions is modeled. By combining both models (GIS informs BN), each considered hospital can be individually assessed with regard to its exposure to the considered disruptions. In this work, we presented a proof of concept of the approach. The preliminary results (see Figure 5) show that hospitals highly exposed to the considered disruptions can quickly be identified as well as hospitals that are not affected at all. Hospitals that are exposed to a subset of the considered disruptions can also be identified, but require a more detailed examination to assess the severity.

Nevertheless, in order to enable a more detailed and realistic assessment of hospitals exposed to flood-induced disruptions, several steps are required within future work to improve the approach. The resolution of the node *Level of Service* could be enhanced, thus considering a more detailed description of hospital services. These services can vary depending on the type and resulting capabilities of a hospital, i.e. some hospitals have specialized services such as fire burn status beds (Fekete et al., 2017). In addition, the resolution to assess the spatial extent of the disruptions can be improved. For example, in the presented work, only a binary assessment of whether a hospital is within the flooded area or not is conducted. Within future work, the height of the flood at each hospital building could be taking into account - data on the flood height is also made available by public services. The method to assess road accessibility could also be improved. In the presented approach, only a single ring buffer is used to determine access roads of a hospital. This analysis neglects the accessibility of these roads from a greater distance - a part of a road near a hospital might not be flooded, but the road might not be accessible from a greater distance. In addition, the type of access roads (e.g. a highway or a small road) could be considered as well as if the access roads to a hospital are passable. The method to assess potential blackout areas can also be improved to allow a higher resolution of blackout areas, e.g. by implementing the weighted Voronoi diagram method presented by Pala et al., 2014 or an analysis based on a reconstruction of the power grid (e.g. see Medjroubi et al., 2017).

When enhancing the complexity of the approach, more data is required to build the BN. Especially, the conditional probability table attached to node *Level of Service* requires extensively more data when considering a higher resolution of both the parent (disruption) nodes and the granularity of the description of hospital services (states of node *Level of Service*). As this data is likely to originate from expert surveys, supporting tools (e.g. see Morris et al., 2014) should be used to elicit probability distributions from experts. However, Bayesian networks also allow data from different sources, such as historical or simulated data, to be combined, which is a major advantage when modeling rare event scenarios, such as the extreme flood scenario of the case study.

To conclude, we presented our work-in-progress approach to assess the effects of multiple (cascading) disruptions on the services of hospitals in a flood scenario. The approach is based on a combination of two models (GIS and Bayesian network) allowing to spatially assess the impact of the considered disruptions on individual hospitals under consideration of uncertainties in disruption exposure and impact. In this work, we presented a proof of concept of the approach in a case study of an extreme flood scenario in the city of Cologne. The preliminary results are promising and enable a first comparative overview of the expected level of services of the examined hospitals. This can support decision-makers in emergency management, but also researchers of other disciplines to better visualize and understand the abstract concepts of cascading effects. It can also help to illustrate how different sectors, such as health, energy, and transport, can be analyzed using one model.

REFERENCES

Ankan, A., & Panda, A. (2015). Pgmpy: Probabilistic graphical models using python. *Proceedings of the 14th Python in Science Conference (SCIPY 2015)*.

Fekete, A. (2019). Critical infrastructure and flood resilience: Cascading effects beyond water. *WIREs Water*, *6*(5), e1370. https://doi.org/10.1002/wat2.1370

Fekete, A. (2020). Critical infrastructure cascading effects. Disaster resilience assessment for floods affecting city of Cologne and Rhein-Erft-Kreis. *Journal of Flood Risk Management*, *13*(2), e312600. https://doi.org/10.1111/jfr3.12600

Fekete, A., Tzavella, K., & Baumhauer, R. (2017). Spatial exposure aspects contributing to vulnerability and resilience assessments of urban critical infrastructure in a flood and blackout context. *Natural Hazards*, *86*(1), 151–176. https://doi.org/10.1007/s11069-016-2720-3

Fink, A., Ulbrich, U., & Engel, H. (1996). Aspects of the January 1995 flood in Germany. *Weather*, *51*(2), 34–39. https://doi.org/10.1002/j.1477-8696.1996.tb06182.x

Geiß, C., Priesmeier, P., Pelizari, P., Calderon, A., Schoepfer, E., Riedlinger, T., Villar-Vega, M., Santa Maria, H., Gomez Zapata, J., Pittore, M., So, E., Fekete, A., & Taubenbo¨ ck, H. (2022). Benefits of global earth observation missions for disaggregation of exposure data and earthquake loss modeling: Evidence from santiago de chile. *Natural Hazards*, *119*, 1–26. https://doi.org/10.1007/s11069-022-05672-6

Herget, J., & Meurs, H. (2010). Reconstructing peak discharges for historic flood levels in the city of Cologne, Germany. *Global and Planetary Change*, *70*(1), 108–116. https://doi.org/10.1016/j.gloplacha.2009.11.011

Johnson, S., Low-Choy, S., & Mengersen, K. (2012). Integrating Bayesian networks and geographic information systems: Good practice examples. *Integrated Environmental Assessment and Management*, *8*(3), 473–479. https://doi.org/10.1002/ieam.262

Jordahl, K., den Bossche, J. V., Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur, A., . . . Leblanc, F. (2020, July). *Geopandas/geopandas: V0.8.1* (Version v0.8.1). Zenodo. https://doi.org/10.5281/zenodo.3946761

Kameshwar, S., Cox, D. T., Barbosa, A. R., Farokhnia, K., Park, H., Alam, M. S., & van de Lindt, J. W. (2019). Probabilistic decision-support framework for community resilience: Incorporating multi-hazards, infrastructure interdependencies, and resilience goals in a Bayesian network. *Reliability Engineering & System Safety*, *191*, 106568. https://doi.org/10.1016/j.ress.2019.106568

Medjroubi, W., Mu¨ ller, U. P., Scharf, M., Matke, C., & Kleinhans, D. (2017). Open Data in Power Grid Modelling: New Approaches Towards Transparent Grid Models. *Energy Reports*, *3*, 14–21. https://doi.org/10.1016/j.egyr.2016.12.001

Melnychuk, E., Sallade, T. D., & Kraus, C. K. (2022). Hospitals as disaster victims: Lessons not learned? *Journal of the American College of Emergency Physicians Open*, *3*(1), e12632. https://doi.org/10.1002/emp2.12632

Mentges, A., Halekotte, L., Schneider, M., Demmer, T., & Lichte, D. (2023). A resilience glossary shaped by context: Reviewing resilience-related terms for critical infrastructures. *International Journal of Disaster Risk Reduction*, *96*, 103893. https://doi.org/10.1016/j.ijdrr.2023.103893

Morris, D. E., Oakley, J. E., & Crowe, J. A. (2014). A web-based tool for eliciting probability distributions from experts. *Environmental Modelling and Software*, *52*, 1–4. https://doi.org/10.1016/j.envsoft.2013.10.010

Nick, F. C., Sa¨nger, N., van der Heijden, S., & Sandholz, S. (2023). Collaboration is key: Exploring the 2021 flood response for critical infrastructures in Germany. *International Journal of Disaster Risk Reduction*, *91*, 103710. https://doi.org/10.1016/j.ijdrr.2023.103710

Pala, O., Wilson, D., Bent, R., Linger, S., & Arnold, J. (2014). Accuracy of Service Area Estimation Methods Used for Critical Infrastructure Recovery. In J. Butts & S. Shenoi (Eds.), *Critical Infrastructure Protection VIII* (pp. 173–191). Springer. https://doi.org/10.1007/978-3-662-45355-1 12

Pearl, J. (1988, September). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.

Puga, J. L., Krzywinski, M., & Altman, N. (2015). Bayesian networks. *Nature Methods*, *12*(9), 799–800. https://doi.org/10.1038/nmeth.3550

Ram´ırez-Agudelo, O. H., Ko¨ pke, C., Guillouet, Y., Scha¨fer-Frey, J., Engler, E., Mielniczek, J., & Sill Torres, F. (2021). An Expert-Driven Probabilistic Assessment of the Safety and Security of Offshore Wind Farms (P. Baraldi, R. Razavi-Far, & E. Zio, Eds.). *Energies*.

Ricci, F., Yang, M., Reniers, G., & Cozzani, V. (2024). Emergency response in cascading scenarios triggered by natural events. *Reliability Engineering & System Safety*, *243*, 109820. https://doi.org/10.1016/j.ress.2023.109820

Rinaldi, S., Peerenboom, J., & Kelly, T. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems Magazine*, *21*(6), 11–25. https://doi.org/10.1109/37.969131

Tzavella, K., Fekete, A., & Fiedrich, F. (2018). Opportunities provided by geographic information systems and volunteered geographic information for a timely emergency response during flood events in cologne, germany. *Natural Hazards*, *91*. https://doi.org/10.1007/s11069-017-3102-1

Valtorta, M., Kim, Y.-G., & Vomlel, J. (2002). Soft evidential update for probabilistic multiagent systems. *International Journal of Approximate Reasoning*, *29*(1), 71–106. https://doi.org/10.1016/S0888-613X(01)00056-1

Witte, D., Bach, S., Lichte, D., Fiedrich, F., & Wolf, K.-D. (2021). Functional impact analysis for complex critical infrastructure systems. *Proceedings of the 31st European Safety and Reliability Conference (ESREL 2021)*. https://doi.org/10.3850/978-981-18-2016-8 280-cd

Homophily in opinion networks affects collective risk perception in heterogeneous populations

Francesca Giardini
Department of Sociology, Faculty of Behavioural and Social Sciences, University of Groningen, The Netherlands
f.giardini@rug.nl

Daniele Vilone
Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy
daniele.vilone@istc.cnr.it

Lorenzo Zino
Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy
lorenzo.zino@polito.it

Ming Cao
Engineering and Technology Institute Groningen, Faculty of Science and Engineering, University of Groningen, The Netherlands
m.cao@rug.nl

ABSTRACT

Understanding how to accurately inform the population about the risks of disasters is key for the well-functioning of our societies. Risk communication is also an essential feature of the disaster cycle, fundamentally contributing to preparedness but also to management. Here, we use an agent-based model (ABM) to investigate such an important problem. Specifically, we study the emergent behavior of a population of individuals who revise their opinion on the risk of a certain event, based on information received from an institution, processed through individual sensitivity, and discussed with peers. Such a complex process may include several biases, e.g., due to heterogeneous risk perception across the population, and homophily, i.e., tendency of individuals to interact with like-minded people. Our ABM, which encapsulates these crucial features, allows us to perform a campaign of numerical simulations towards gaining mathematically-grounded insights into their impact on the emergent behavior of the population and, ultimately, on how accurately institutional information is received and processed by a population. Such insights can be useful to design empirical studies to test them and, in case of empirical support, to use them to design recommendations for policy decision makers.

Keywords

Agent-Based Model, Homophily, Opinion Dynamics, Risk Perception, Network

INTRODUCTION

Accurately informing the population about the risks connected to natural hazards, climate change, and men-made disasters is crucial. The importance of sharing accurate and useful information about risks is growing, together with the increase in the number of disaster events all over the world (*Global Assessment Report On Disaster Risk Reduction*, 2022). Effective risk communication means that citizens are correctly informed about what might happen, how they should prepare and what they should do to prevent, mitigate or adapt to the consequences. In a way, risk communication is crucial for response and management (Mamuji & Etkin, 2019), but its effects are present also in the recovery and mitigation phase. However, effective risk communication is extremely hard to achieve, and its effectiveness can depend on several factors (Bradley et al., 2014).

One major challenge comes from the fact that the same message can be interpreted in different ways, but also that the process of sharing information can lead to a final outcome, in terms of content and its spread, that is impossible

to predict. The social amplification of risk framework (SARF), first proposed by Kasperson et al., 1988, was intended as a way to explain how risks can be amplified in social interactions. Its purpose is to show the multiple interactions between psychological, social and cultural factors that can heighten or attenuate public perceptions of risk and related risk behaviours. The SARF has been applied to many different situations and hazards, ranging from wildfire risks (Brenkert-Smith et al., 2012) to chemical accidents (de Souza Porto & de Freitas, 1996), and to nuclear energy generation in Europe (Bearth & Siegrist, 2021). However, several decades after the introduction of the SARF, there are open issues that are still puzzling researchers and practitioners (Kasperson et al., 2022): social media, uncertainty, risk communication and integrated analysis. In this work, we focus on risk communication and systemic risk (which is referred to as integrated analysis by (Kasperson et al., 2022)) and we aim to contribute to understanding how collective risk perception emerges from artificial agents' beliefs, but also how it changes and evolves over time.

Agent-based modeling offers the opportunity to explore the factors impacting the spread of reassuring or worrying messages (Onggo et al., 2014)), then contributing to identify possible pathways for the social amplification of risk (Busby et al., 2016). A parallel but unrelated strand of research about the spread of beliefs in a population is opinion dynamics, a well-known sub-field of complex systems and agent-based modeling which aims to model the determinants of opinion spreading in a population, and predict its emergent behavior, which can span from the emergence of consensus among individuals, to steady disagreement andpolarization (Deffuant et al., 2000; Hegselmann & Krause, 2002; Proskurnikov & Tempo, 2017; Sen & Chakrabarti, 2014).

An early attempt to combine an opinion-dynamical model with a model of collective risk perception (Giardini & Vilone, 2021) has provided interesting insights about the impact of trust in institutional communication, trust in peers and varying level of agents' individual risk sensitivity. This model of collective risk perception based on opinion dynamics shows that, even if the agents receive the same initial information about the likelihood of the hazard, their perception of the actual risks is modified by individual traits and by the resulting social influence process. For different parameter configurations, overestimating risks (with regard to the initial state and the information received) is a likely outcome, but the system never reaches full consensus. This suggests that, regardless of the content of the message, individual risk perception and belief revision processes will lead the population to hold different beliefs about how risky a certain situation is. Even if the message has the same content, heterogeneous agents will perceive and transmit it according to their individual features, meaning that collective risk perception can be feebly linked to the original message. Another interesting but rather counter-intuitive result of that model is that different network structures (including regular structure, random topologies, and real-world structures) had negligible effects on the final opinion distribution, even on two empirically calibrated network topologies. This result is at odds with empirical studies (Becker et al., 2017) and simulation models of opinion spreading on networks (Flache et al., 2017), which suggest that network structure plays an important role in opinion formation and change. In order to better understand the interplay between network structure and collective risk perception, we designed a model in which information about risk is spread on the basis of the similarity across agents, i.e., their homophily. Homophily, *the observed tendency of like to associate with like* (Kossinets & Watts, 2009), is one of the most recurrent empirical regularities of social life (McPherson et al., 2001). Friends, romantic partners, colleagues, and other kinds of associates all tend to be more similar to each other than randomly chosen members of the same population with respect to a variety of dimensions (race, age, gender, socioeconomic status, habits, education, hobbies). Homophily can foster the spread of healthy behaviors in social networks (Nunner et al., 2022), but it can also contribute to the formation of echo chambers and polarization (Bessi et al., 2016). It is worth noting that, to the best of our knowledge, homophily in disaster risk perception has not been previously addressed. In the risk and decision making literature some attention has been devoted to risks related to different health treatments (Berry et al., 2018), and infectious disease outbreak (Kadelka & McCombs, 2021), but we are not aware of any work focusing on the relationship between homophily in risk perception and risk communication.

In order to fill in this gap, we explore the effect of homophily-based communication in an opinion dynamic model of collective risk perception. Building upon the work of (Giardini & Vilone, 2021), we designed an agent-based model (ABM) in which agents receive information about the likelihood of a disaster from an institutional source (broadcast), revise this information according to their own individual traits (risk perception), and then share the information with other agents (risk communication) on the basis of their homophily.In our model, agents interact in pairs, thus resembling an interpersonal discussion on risk, which is an important means through which individuals gather information and make sense of uncertain situations (Perlstein, 2023). Binder and colleagues in their study (Binder et al., 2010) using data from a public opinion survey on a new and potentially risky facility in the US found empirical support for the hypothesis that people are more likely to engage in interpersonal conversations with interaction partners who are expected to share their 'discussion valence', i.e., like minded individuals. This is in line with studies on political communication that reported that it is relatively common for people to find themselves in discussion networks with people sharing similar opinions (Huckfeldt & Sprague, 1987). Homophilous ties can

also be established on the basis of other characteristics, for instance shared political views or trust in institutions, but in this study we are interested in understanding whether and how preferentially interacting with others who have similar levels of risk perception might lead to consensus on risk perception in the population. In an extension of the model (in preparation) we plan to add homophily in trust in institutions and trust in the media as additional conditions to explore whether being homophilous in different dimensions has an effect on the final distribution of risk perception in the population.

Inspired by similar approaches in mathematical modeling of epidemic spreading (Frieswijk et al., 2023), we encapsulate homophily into the risk communication step of the ABM. Specifically, we include an additional parameter that captures the level of homophily of the population, which shapes the network formation process through which agents share information with others, favoring interactions among agents with the same risk sensitivity. Ultimately, our goal is to employ the ABM to understand whether homophily leads to highly clustered and polarized risk perceptions, and whether this varies depending on the levels of perceived risk.

MODEL DESCRIPTION

In our ABM, we consider a population of L agents, indexed by positive integer numbers $\mathrm{V} = \{1, \ldots, L\}$. Each agent $i \in \mathrm{V}$ is characterized by an opinion $O_i \in [0, 1]$, which represents their evaluation of the risk. We shall refer to this variable as the *opinion* of agent i, consistent with the opinion dynamics literature. Specifically, $O_i = 0$ represents the scenario in which agent i's evaluation is of minimum risk, whereas $O_i = 1$ means that i gives the maximum evaluation of the risk. For the sake of simplicity, we model opinions as the subjective probability that the disaster will actually take place, regardless of the magnitude of the possible consequences. We acknowledge that this is an important issue and we will tackle this in future work. Agents' opinions are updated on the bases of the interplay between internal characteristics of the agents and two different sources of influence: i) elaboration of institutional information, and ii) interaction among peers.

The elementary dynamics is set as follows: at the beginning of each discrete time step, the institution communicates to each and every agent the value of the risk I, which is a continuous value in the range $I \in [0, 1]$ with the same convention that $I = 0$ means minimum risk and $I = 1$ means maximum risk. Here, we assume that the institutional information I is kept constant throughout the iterations. The agents elaborate the information received. Then, they share their updated opinions among each other, and further elaborate their opinion on the base of such interaction with peers. The mathematical details of such opinion revision process are described in the following section.

Elaboration of Institutional Information

Once received the value I of risk assessment by the institution, each agent i modifies their opinion on risk o_i according to the Deffuant's model (Deffuant et al., 2000). Specifically, the current opinion (O_i^o) is revised to the new opinion (O_i) as

$$O_i^o \longrightarrow O_i = O_i^o + T_i (I - O_i^o) , \tag{1}$$

where $T_i \in [0, 1]$ is a parameter that captures individual's trust in the institution: the larger such parameter, the larger the impact of the information provided by the institution in agent's i opinion formation process.

Then, the updated opinion O_i is further processed according to internal characteristics of the agent. Specifically, we introduce a parameter $R_i \in \{-1, 0, +1\}$ that captures i's risk sensitivity. When $R_1 = -1$, agent i has low risk sensitivity, and is thus inclined to underestimate the risk; when $R_1 = +1$, agent i is highly sensitive to risk, tending to overestimate it; while $R_i = 0$ captures risk-neutral agents who do not distort the information received. Such a further process based on risk sensitivity is captured by the following update rule:

$$O_i \longrightarrow \begin{cases} \frac{1}{2}(1 + O_i) & \text{if} \quad R_i = +1 \\ O_i & \text{if} \quad R_i = 0 \\ \frac{1}{2} O_i & \text{if} \quad R_i = -1 . \end{cases} \tag{2}$$

In plain words, the presence of (positive or negative) risk sensitivity induces a bias affecting how individuals process the information received from the institution. This is consistent with what observed in the social psychology

literature (see, e.g., Popovic et al., 2020; Richard Eiser et al., 2012).

Table 1. Variables and Parameters of the Model

Symbol	*Values*	*Meaning*
L	Integer	Number of agents in the network
O_i	$[0, 1]$	Evaluation of the risk (opinion) of agent i
I	$[0, 1]$	Risk assessment communicated by the institution
T_i	$[0, 1]$	Trust in the institution of agent i
R_i	$\{-1, 0, +1\}$	Risk sensitivity of agent i
P_i	$[0, 1]$	Trust in peers of agent i
H	$[0, 1]$	Level of homophily in the network

Interaction Among Peers

Once the agents have received and processed the institutional information I, they exchange their opinions with peers on a social influence network, consistent with Scherer and Cho, 2003. More precisely, for each time step, we pick up L couples of agents (i, j), sampled uniformly at random. Each of these couples is a candidate interaction that can take place. Once the pair (i, j) is selected, we randomly assign an order to the two agents that form the edge, say i and j. Let us define j as the "speaker" and i as the "listener" (the symmetrical interaction where i is the speaker and j the listener will take place in the same way).

At this stage, we incorporate homophily in the model as the possibility that individuals may prefer to interact with akin peers than strangers. This assumption is based on the homophily principle, which means that we assume that agents with the same risk sensitivity interact more often (McPherson et al., 2001). Homophily is modeled as follows. For each one of the L couples of agents (i, j) sampled at random, i and j actually interact according to a probabilistic rule. Specifically, the interaction always takes place if the two agents i and j have the same risk sensitivity. Otherwise, the interactions takes place only with probability equal to $1 - H$, while with probability equal to H the "listener" refuses to receive information from the sender, and thus link (i, j) is discarded. In other words, if $H = 0$ there is no homophily in the population, and everyone interacts indifferently with everyone else; at the opposite of the spectrum, if $H = 1$, agents avoid completely interactions with peers not belonging to the same class. Clearly, such a mechanism ultimately generates a time-varying network of social contacts, whose pattern is shaped by homophily. In fact, for $H = 0$, at each discrete time step individuals interacts on an instance of an Erdos-Renyi random graph with L links (Newman, 2010); as H increases, the network starts displaying a clustered structure, which will ultimately impact the emergent behavior of the system, as we shall see in the following section.

Finally, if the interaction between a speaker j and a listener i takes place, then the listener revises their opinion. Specifically, let O_i and O_j be their opinions before such interaction takes place, respectively. Then, agent i will change their own opinion according to the Deffuant's rule described in Deffuant et al., 2000:

$$O_i \longrightarrow O_i' = O_i + P_i (O_j - O_i) , \tag{3}$$

where $P_i \in [0, 1]$ is a parameter that captures agent i's trust in peers. In this work, we set $P_i \equiv 1 - T_i$, which yields

$$O_i \longrightarrow O_i' = O_i + (1 - T_i)(O_j - O_i) . \tag{4}$$

On the other hand, the speaker does not revise their opinion. Finally, the listener elaborates again its opinion according to the rule in Eq. (2). All the variables and parameters of the model are summarized in Table 1.

Simulation Setting

In this work, we perform a campaign of numerical simulation on the model described in the above, in order to explore the role of homophily on the collective risk perception. To this aim, we generate Monte Carlo simulations of the ABM, and we estimate the quantity of interest (i.e., the final average opinion of the population), by averaging the result obtained over an ensemble of 2,000 independent realizations (*i.e., runs*).

In each run, we consider $L = 1,000$ agents and we initialize agents' initial opinions and trust in institution uniformly at random in the domain $[0, 1]$, each agent independently of the others. The risk sensitivity is assigned as follows: each agent has a probability $p = 0.5$ to have risk sensitivity equal to zero (neutral agent), a probability $\rho/2$ to have $R_i = +1$ (risk-averse agent) and $(1 - \rho)/2$ to have $R_i = -1$ (risk-prone agent), where $\rho \in [0, 1]$ captures the average bias of the population: $\rho < 0.5$ means that there are more risk-prone agent than risk-averse agent, while $\rho > 0.5$ represents the opposite scenario.

Every run is performed for a sufficiently large number of iterations in order to achieve a final steady state, *i.e.*, a configuration where the dynamics has become constant and the global configuration of the system is stable, that is, the state of the system does not change anymore.

RESULTS

In this section, we describe the results of our numerical simulations. Specifically, we perform two sets of numerical simulations. In the first one, we consider the scenario with no homophily ($H = 0$). In this scenario, we consider populations with different characteristics in terms of risk sensitivity (captured by the parameter ρ). Specifically, we consider a balanced scenario ($\rho = 0.5$), and two biased scenarios (with $\rho = 0.15$ and $\rho = 0.85$, respectively). In all scenarios, we consider different values of the risk assessment communicated by the institution I, and we study how the opinion revision process shapes the collective risk perception.

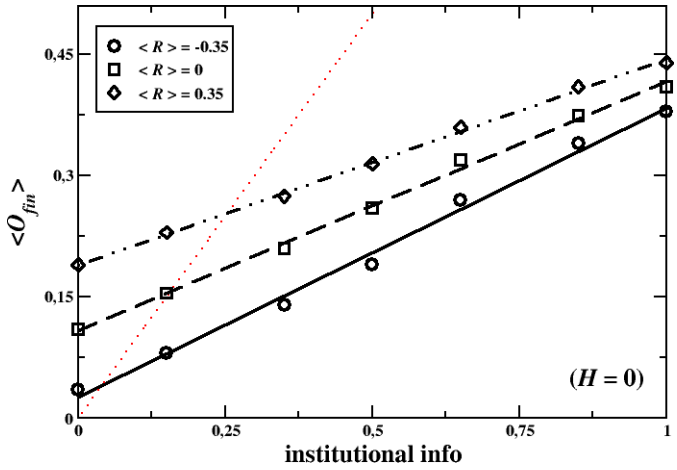


Figure 1. Monte Carlo estimation of the average final opinion $\langle O_{fin} \rangle$ as a function of the institutional information I for a system of $L = 1,000$ agents and a population with initial opinions and trust towards institution randomly assigned with uniform distributions, for different values of risk sensitivity. Lines represent linear fits of numerical data, red-dotted one is the bisector of the first orthant.

Specifically, Figure 1 shows the final average opinion, that is, the average evaluation of the risk in the population, in a completely balanced system, that is, when $\rho = 0.5$, yielding a balance between the number of agents with high and low risk sensitivity (on average). In general, our results are consistent with those described in Giardini and Vilone, 2021: the population tends on average to overestimate the risk for low-risk situations (that is, small values of I), and underestimate it in the opposite case. Interestingly, the value of I for which the population starts to underestimate the risk is lower than $I = 0.5$. The same study is reproduced also in the case of non-balanced populations, that is, for populations with lower and higher risk sensitivity. Overall, the emergent behavior is qualitatively similar. Unsurprisingly, the collective risk perception is higher for the population that is more sensitive to risk. However, it is worth noticing that even in the scenario in which $\rho = 0.85$ (yielding an average risk sensitivity of 0.35), the population starts underestimating the risk even for values of institutional information lower than $I = 0.5$. These observations are consistent with the conclusions of previous research, which highlighted how the nontrivial way through which people elaborate information received from the institution may lead to a biased collective risk perception; in particular, with an underestimation of severe risks (Giardini & Vilone, 2021).

In the second set of simulations, whose results are illustrated in Figure 2, we consider a scenario in which homophily is present. Specifically, we investigate how increasing the value of homophily shapes the final average opinion of the population. To this aim, we consider the basis balanced scenario presented in Figure 1 (with $\rho = 0$), and we estimate the average final opinion for different values on H. Our results suggest that increasing the homophily can lead to an overestimation of the risk, whereby the echo chambers formed by the network formation process seem to amplify the risk perception, with potentially harmful consequences to the society.

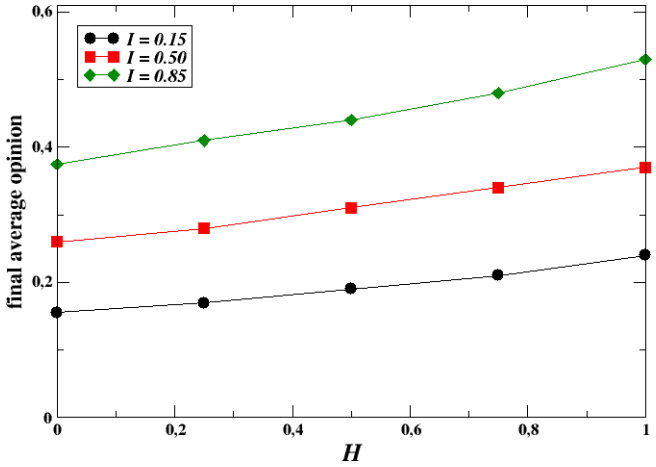


Figure 2. Monte Carlo estimation of the average final opinion $\langle O_{fin} \rangle$ as a function of the homophily H for different values of the institutional information. The system has $L = 1,000$ agents with fully balanced parameters: initial opinions, trust towards institution, and risk sensitivity randomly assigned with uniform distributions.

CONCLUSION AND ONGOING RESEARCH

In this work, we used agent-based modeling to explore the role of homophily on the collective risk perception in heterogeneous populations. Specifically, building on an ABM proposed in Giardini and Vilone, 2021, we have incorporated a mechanism that captures the tendency of individuals to interact more with like-minded people. Then, we have performed a campaigns of Monte Carlo simulations to unveil the impact of such behavioral mechanism on the emergent behavior of the system, showing how it could lead to biases in how institutional information is received and processed by a population.

The results discussed in the previous section offer us a preliminary understanding of the interplay between risk communication, individual biases and homophily in information spreading. When agents interact preferentially with similar agents this results in a widespread amplification of the risk perception, which can have several undesired outcomes. This finding is in line with research on homophily in the spreading of desirable and undesirable behaviors (Brechwald & Prinstein, 2011), paving the way for different lines of further research toward extending the existing findings, deriving analytical results to support them, and validating the model on real-worlds case studies. We are aware of the multiple limitations of this work in progress, including lack of validation of the model, limited exploration of the parameter space and a quite simplistic description of the homophily principle.

To address some of these limitations, we are currently working along three main avenues in order to further develop the model.

First, the numerical simulations presented in this paper are being extended along several directions. One relevant direction is to embed the model on a nontrivial network structure, restricting social interactions to those between neighbors on the network. This approach, already used in Giardini and Vilone, 2021 for the simpler model without homophily, can be readily extended to incorporate homophily toward investigating the interplay between such behavioral mechanisms and the complexity of real-world networks of social interactions. Moreover, more realistic implementations of homophily should be considered and incorporated into the model.

A second research direction that we are currently exploring is the possibility to establish rigorous analytical results for the model of collective risk perception formation described in this paper. Concerning this research direction, the key step consists in re-formulating the opinion formation process described in Eqs. (1)–(4) using the system-theoretic formalism of opinion dynamics models [See, e.g., Proskurnikov and Tempo, 2017] Using this formalism, under some assumptions, it is possible to gain analytical insights into the emergent behavior of the population, and derive closed-form expressions for the final average opinion, as function of the different parameters of interest. Some promising preliminary steps towards achieving such a goal can be found in Zino et al., 2024. This procedure will allow us to corroborate the findings of our numerical simulations with rigorous analytical observations.

Third, we are working towards validating the ABM in real-world scenarios, using survey and experimental data to calibrate the model. A recent review paper (Bradley et al., 2014) pointed out that natural disaster preparedness interventions, which authorities might wish to target at a whole population, mainly involved interpersonal communication, but their effectiveness is limited. The insights gathered thanks to ABM can be used to shape theories and data collection on the factors driving an heterogeneous population to form its collective risk perception about a certain event. This can be done starting from survey data, such as the study in Salvati et al., 2014, which collects data on risk perception concerning floods and landslides in different Italian regions, and how such collective opinion evolves over time. Using this data, we are currently working towards calibrating the model to provide an empirical validation that supports our ABM. We also plan to extend the model in order to design opinions about hazards differing in terms of impact.

REFERENCES

Bearth, A., & Siegrist, M. (2021). The social amplification of risk framework: A normative perspective on trust? *Risk Analysis*, *42*(7), 1381–1392. https://doi.org/10.1111/risa.13757

Becker, J., Brackbill, D., & Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences*, *114*(26), E5070–E5076. https://doi.org/10.1073/pnas.1615978114

Berry, D. L., Blonquist, T. M., Pozzar, R., & Nayak, M. M. (2018). Understanding health decision making: An exploration of homophily. *Social Science & Medicine*, *214*, 118–124. https://doi.org/10.1016/j.socscimed.2018.08.026

Bessi, A., Petroni, F., Vicario, M. D., Zollo, F., Anagnostopoulos, A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Homophily and polarization in the age of misinformation. *The European Physical Journal Special Topics*, *225*(10), 2047–59. https://doi.org/10.1140/epjst/e2015-50319-0

Binder, A. R., Scheufele, D. A., Brossard, D., & Gunther, A. C. (2010). Interpersonal amplification of risk? citizen discussions and their impact on perceptions of risks and benefits of a biological research facility. *Risk Analysis*, *31*(2), 324–334. https://doi.org/10.1111/j.1539-6924.2010.01516.x

Bradley, D. T., McFarland, M., & Clarke, M. (2014). The effectiveness of disaster risk communication: A systematic review of intervention studies. *PLoS Currents*. https : / / doi . org / 10 . 1371 / currents . dis . 349062e0db1048bb9fc3a3fa67d8a4f8

Brechwald, W. A., & Prinstein, M. J. (2011). Beyond homophily: A decade of advances in understanding peer influence processes. *Journal of Research on Adolescence*, *21*(1), 166–179. https://doi.org/10.1111/j.1532-7795.2010.00721.x

Brenkert-Smith, H., Dickinson, K. L., Champ, P. A., & Flores, N. (2012). Social amplification of wildfire risk: The role of social interactions and information sources. *Risk Analysis*, *33*(5), 800–817. https://doi.org/10.1111/j.1539-6924.2012.01917.x

Busby, J., Onggo, B., & Liu, Y. (2016). Agent-based computational modelling of social risk responses. *European Journal of Operational Research*, *251*(3), 1029–1042. https://doi.org/10.1016/j.ejor.2015.12.034

Deffuant, G., Neau, D., Amblard, F., & Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, *03*(01n04), 87–98. https://doi.org/10.1142/s0219525900000078

de Souza Porto, M. F., & de Freitas, C. M. (1996). Major chemical accidents in industrializing countries: The socio-political amplification of risk. *Risk Analysis*, *16*(1), 19–29. https://doi.org/10.1111/j.1539-6924.1996.tb01433.x

Flache, A., Ma¨s, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., & Lorenz, J. (2017). Models of social influence: Towards the next frontiers. *Journal of Artificial Societies and Social Simulation*, *20*(4). https://doi.org/10.18564/jasss.3521

Frieswijk, K., Zino, L., & Cao, M. (2023). A polarized temporal network model to study the spread of recurrent epidemic diseases in a partially vaccinated population. *IEEE Transactions on Network Science and Engineering*, 1–12. https://doi.org/10.1109/tnse.2023.3272472

Giardini, F., & Vilone, D. (2021). Opinion dynamics and collective risk perception: An agent-based model of institutional and media communication about disasters. *Journal of Artificial Societies and Social Simulation*, *24*(1), 4. https://doi.org/10.18564/jasss.4479

Global Assessment Report On Disaster Risk Reduction (tech. rep.). (2022). United Nations Office for Disaster Risk Reduction (UNDRR).

Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, *5*(3).

Huckfeldt, R., & Sprague, J. (1987). Networks in context: The social flow of political information. *American Political Science Review*, *81*(4), 1197–1216. https://doi.org/10.2307/1962585

Kadelka, C., & McCombs, A. (2021). Effect of homophily and correlation of beliefs on covid-19 and general infectious disease outbreaks. *PLOS ONE*, *16*(12), 1–20. https://doi.org/10.1371/journal.pone.0260973

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., & Ratick, S. (1988). The social amplification of risk: A conceptual framework. *Risk Analysis*, *8*(2), 177–187. https://doi.org/10.1111/j.1539-6924.1988.tb01168.x

Kasperson, R. E., Webler, T., Ram, B., & Sutton, J. (2022). The social amplification of risk framework: New perspectives. *Risk Analysis*, *42*(7), 1367–1380. https://doi.org/https://doi.org/10.1111/risa.13926

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, *115*(2), 405–450. https://doi.org/10.1086/599247

Mamuji, A. A., & Etkin, D. (2019). Disaster risk analysis part 2: The systemic underestimation of risk. *Journal of Homeland Security and Emergency Management*, *16*(1), 20170006. https://doi.org/10.1515/jhsem-2017-0006

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Newman, M. (2010). *Networks: An introduction*. Oxford University Press.

Nunner, H., Buskens, V., Teslya, A., & Kretzschmar, M. (2022). Health behavior homophily can mitigate the spread of infectious diseases in small-world networks. *Social Science & Medicine*, *312*, 115350. https://doi.org/10.1016/j.socscimed.2022.115350

Onggo, B. S., Busby, J., & Liu, Y. (2014). Using agent-based simulation to analyse the effect of broadcast and narrowcast on public perception: A case in social risk amplification. *Proceedings of the Winter Simulation Conference 2014*, 322–333. https://doi.org/10.1109/WSC.2014.7019899

Perlstein, S. (2023). Risk perception and interpersonal discussion on risk: A systematic literature review. *Risk Analysis*. https://doi.org/10.1111/risa.14264

Popovic, N. F., Bentele, U. U., Pruessner, J. C., Moussa¨ıd, M., & Gaissmaier, W. (2020). Acute stress reduces the social amplification of risk perception. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-62399-9

Proskurnikov, A. V., & Tempo, R. (2017). A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control*, *43*, 65–79. https://doi.org/10.1016/j.arcontrol.2017.03.002

Richard Eiser, J., Bostrom, A., Burton, I., Johnston, D. M., McClure, J., Paton, D., van der Pligt, J., & White, M. P. (2012). Risk interpretation and action: A conceptual framework for responses to natural hazards. *International Journal of Disaster Risk Reduction*, *1*, 5–16. https://doi.org/10.1016/j.ijdrr.2012.05.002

Salvati, P., Bianchi, C., Fiorucci, F., Giostrella, P., Marchesini, I., & Guzzetti, F. (2014). Perception of flood and landslide risk in italy: A preliminary analysis. *Natural Hazards and Earth System Sciences*, *14*(9), 2589–2603. https://doi.org/10.5194/nhess-14-2589-2014

Scherer, C. W., & Cho, H. (2003). A social network contagion theory of risk perception. *Risk Analysis*, *23*(2), 261–267. https://doi.org/10.1111/1539-6924.00306

Sen, P., & Chakrabarti, B. K. (2014). *Sociophysics: An Introduction* (1st ed.). OUP Oxford.

Zino, L., Giardini, F., Vilone, D., & Cao, M. (2024). On modeling collective risk perception via opinion dynamics. To appear in the *Proceedings of the 2024 European Control Conference*.

Sensitivity Analysis of a Road Clearing and Relief Supplies Distribution Model for Vancouver Island in the Event of a Cascadia Megathrust Earthquake

Ronald Pelot
Industrial Engineering
Dalhousie University, Halifax, Canada
Ronald.Pelot@Dal.ca

Pranitha Vattoni
Industrial Engineering
Dalhousie University, Halifax, Canada
Pranitha.Vattoni@Dal.ca

Luana Souza Almeida
Industrial Engineering
Dalhousie University, Halifax, Canada
Luana.Almeida@Dal.ca

Floris Goerlandt
Industrial Engineering
Dalhousie University, Halifax, Canada
Floris.Goerlandt@Dal.ca

ABSTRACT

Southwestern British Columbia is the most seismically active region in Canada because of its proximity to the Cascadia Subduction zone. To improve emergency preparedness for Vancouver Island, models have previously been developed to estimate the potential damage to infrastructure from a major earthquake, to consider community resilience while awaiting relief supplies, and to optimize procedures for repairing damaged roads and routing trucks from marine terminals to needy communities during the 72-hour emergency response phase. However there are large uncertainties in the disaster impacts, and numerous assumptions in the models' parameters, so this study serves to conduct sensitivity analysis on several of the factors to produce more robust recommendations from the model outcomes including: best depot location for road-clearing teams; most critical roads for supply distribution; most critical ports for supply distribution; and the effect of incorporating a community resilience measure. Effective display of results is an important consideration.

Keywords

Relief supplies distribution, road-clearing model, earthquake, sensitivity analyses.

INTRODUCTION

The Cascadia Subduction zone is the tectonic boundary between the Continental North American and the Oceanic Juan De Fuca plates in the Pacific Ocean. Southern British Columbia, Canada, lies on top of the Cascadia Subduction zone, making it the most seismically active region in all of Canada, with a 2-month sample shown in Figure 1 according to the (logarithmic) Moment Magnitude Scale (Rogers, 1998).

The closest densely populated regions to the subduction region, the communities on Vancouver Island (VI) and the smaller nearby Gulf Islands, are the most vulnerable places to large magnitude earthquakes. A Cascadia earthquake of magnitude M9, which recurs at intervals from 100 to 1000 years (Clague, 2002), can cause major infrastructure damage on these islands, impeding relief supply distribution during the immediate response phase following the shock, compounded by potential disruption in transportation modes which could have a significant impact on the flow of essential goods between and within the islands.

WiP Paper – Analytical Modeling and Simulation
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

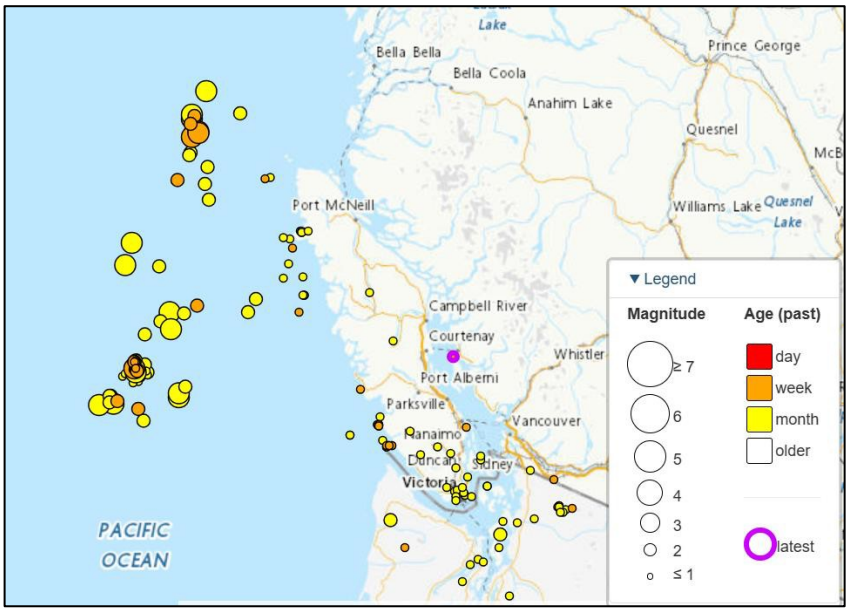


Figure 1. All earthquakes for May and June 2023 (source: Natural Resources Canada, 2010)

A significant contribution to emergency preparedness is to model the main components of the distribution network for relief supplies, how elements may be damaged by a natural disaster, and what mitigating strategies may be considered to minimize the disruption. Studies have been conducted to predict the damage a Cascadia Earthquake may induce on and around Vancouver Island (VI), as well as road network reconstruction and relief supplies distribution plans to deal with the aftermath (Chang et al., 2020). Other models for reconstruction of roads and routing of relief supplies were developed and executed on data available from past disasters (Souza Almeida et al., 2022a). However, little research is available on sensitivity analyses of such models, whether on existing case studies or simulated disaster data. This study deals with analyzing the results of one such model for a simulated Cascadia earthquake scenario on VI.

The relief supplies delivery problem involves boxes of essential goods, with one box per person in the affected region lasting one week. The mechanism is to have the boxes delivered to the islands via ship from the City of Vancouver on the mainland, while having road clearing teams repair roads on VI in the most effective way to reconnect communities whose links to the road network were severed by the earthquake. Delivery trucks pick up the boxes from marine delivery points, and use any undamaged roads or repaired roads to access the communities. However, the earthquake can also shut down ports on the islands temporarily due to damage, and/or halt some shipping routes due to interrupted navigation services. Barges can be used as surge capacity to get boxes to Vancouver Island as they can access many more suitable landing spots. Within the 72-hour emergency phase planning horizon, any unmet demand can be fulfilled by helicopter from mainland airports to an island community.

The post-earthquake optimal resupply plan relies on several foundational models. The Critical Infrastructure (CI) Analysis of Regional Disruption & Ferry Route Interruption model (Bell & Bristow, 2020) characterizes the damages to infrastructure, marine and road transport, and shipping routes given a Cascadia earthquake. Community demand for relief supply boxes is principally proportional to their respective populations, but this can also be weighted by a resilience index (Chang and Tanner, 2022) which captures the relative impact of the earthquake on each community based on multiple factors associated with their preparedness and vulnerability. The resulting community impact factor (CIF) is an integer measure of a community's resilience to the Cascadia earthquake with 1 being the most resilient value and 5 being the least resilient community to an earthquake.

Outputs from those two models concerning CI and CIF respectively produce inputs for relief supply distribution optimization, which comprises two models: a road clearing model on VI to reconnect isolated communities, and a multi-modal distribution model of relief supplies which takes advantage of these reconnected roads to reach affected communities. These two models are synchronised and referred to as the Road Clearing and Relief Supplies Distribution (RCRSD) model. The RCRSD is solved using a Greedy Randomised Adaptive Search Procedure (GRASP) heuristic to optimally satisfy the demand for relief supply boxes as shown in Figure 2. The RCRSD model generates a solution for damaged roads' reconstruction, while simultaneously finding the optimal way to distribute supplies via sea, land, and air to the affected communities.

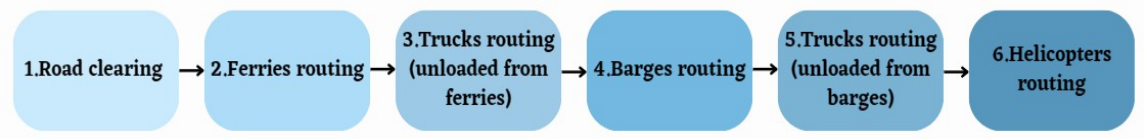


Figure 2. Working order of the RCRSD model

The combination of all of these models provides scenario results to help develop response plans by the relevant authorities, as well as industry and community stakeholders to improve the emergency preparedness for VI. However, there are large uncertainties and multiple assumptions in all stages of each of these models, so the outputs from a single run do not well represent the potential range of outcomes following an actual earthquake. But conversely, it is infeasible to run the models for all possible combinations of conditions, therefore a systematic sensitivity analysis process is needed to reflect the span of post-disaster circumstances and associated risk mitigation strategies to best prepare for the emergency response phase. Such a sensitivity analysis based on the RCRSD model, but also with some variations in outputs from the CI and CIF models, is the subject of this study.

The aim of this paper is to conduct sensitivity analyses on multiple input parameters to the RCRSD model to answer the following questions. The sensitivity analyses are primarily conducted given a situation where the earthquake damage to roads and ports is moderate (Case A: *Partial disconnection*), with an extension considering limited variations for the situation where damage is much greater (Case B: *Extensive disconnection*).

- Where are the best and worst locations to establish the depot on Vancouver Island where the road clearing teams are pre-positioned?
- Which damaged road segments are consistently cleared over multiple runs of the RCRSD model (as the solution method is heuristics based), and how does that affect communities supplied?
- For a given level of roads consistently cleared within the 72-hour window (i.e. in 25%, 50% and 70% of the runs respectively), if those road segments are then considered undamaged, what is the next priority set of roads to be cleared?
- Which ferry port and which barge port respectively supply the most relief boxes, and what is the impact of having one or the other unavailable?
- What is the impact of having the number of available ferries or barges from the mainland reduced?
- What is the effect on the relief box supply distribution if the communities' relative neediness (CIF – community impact factor) is omitted?
- How should the results of the above analyses be presented to be intelligible and useful?

METHODOLOGY

The methodology for conducting the sensitivity analysis is essentially these four major steps, each of which is expanded upon below:

a) Determine which variables will be considered to be altered for the sensitivity analysis;

b) Determine how to represent the results effectively, given the large number of outputs;

c) Determine a framework for altering combinations of the input variables for the sensitivity analysis;

d) From the plethora of results, synthesize recommendations for Emergency Management planners.

Input Variables

Over 50 input variables are used in the RCRSD model. Many of them are considered fixed, such as the road network on Vancouver Island, and the communities' populations. The sensitivity analyses focus on factors within the planners' control, such as where to pre-position road clearing teams, and anticipation of various state variables affected by the earthquake such as which ports will be opened or closed and the number of ferries or barges available.

Output Parameters

For each run of the RCRSD model, 17 output files are generated as shown in Table 1.

Table 1. List of all output files generated from a single run of the RCRSD model

	Output name	Output ID	Content details
1	Initial demand	ID	List of demands of every community on VI in terms of number of boxes of supplies
2	Road clearing edges	RCE	Objective value of road clearing sub-model and list of unblocked edges with the time taken and ID of the team that clears it
3	Road clearing routes	RCR	Route taken by each road clearing team
4	Ferries island delivery	FID	Demands delivered on ferries to each of the smaller islands in terms of number of trucks
5	Ferries demands	FD	Objective value of the ferries routing optimization model and the demands delivered to each port via ferries in terms of number of trucks
6	Ferries routes	FR	Routes details for every ferry
7	Ferries truck delivery	FTD	Demands delivered by trucks that arrived on VI via ferries to each community in terms of number of boxes
8	Ferries trucks routes	FTR	Routes details of trucks that arrived on VI via ferries
9	Barges island delivery	BID	Demands delivered on barges to each of the smaller islands in terms of number of trucks
10	Barges demands	BD	Objective value of the barges routing optimization model and the demands delivered to each port via barges in terms of number of trucks
11	Barges routes	BR	Routes details for every barge
12	Barges truck delivery	BTD	Demands delivered by trucks that arrived on VI via barges to each community in terms of number of boxes
13	Barges trucks routes	BTR	Routes details of trucks that arrived on VI via barges
14	Helicopter community demands	HD	Demands delivered by helicopters to communities on VI
15	Helicopter routes	HR	Routes details of helicopters
16	Communities with unsatisfied demands	UC	List of communities whose demands are not 100% met
17	Communities completely supplied	SC	List of communities whose demands are completely met

Regional Division of Vancouver Island

Given the large number of output variables (Table 1), coupled with the 109 communities on VI included in the model (Figure 3), it is ineffectual to provide the results solely in terms of distinct community supply satisfied, segments of roads fixed within the time horizon (out of 3707 possible), ports and barges and ships operating, etc. While all of these details are calculated and stored, they must be aggregated in effective ways to provide useful insights regarding the outputs.

Various techniques were used to aggregate the results and render them digestible. For some, percentages were applied such as the percent of damaged roads repaired within the specified time horizon, or the percent of total community demand for boxes met via water and land avenues (versus helicopter). To represent changes in supplied amounts under different run conditions, the levels of supplies are often aggregated by four different regions of VI namely North Island, Central Island, Pacific Rim, and South Island as shown in Figure 3. Sunshine Coast and Gulf Islands, although used in the RCRSD model, were excluded from the sensitivity analyses done in this study.

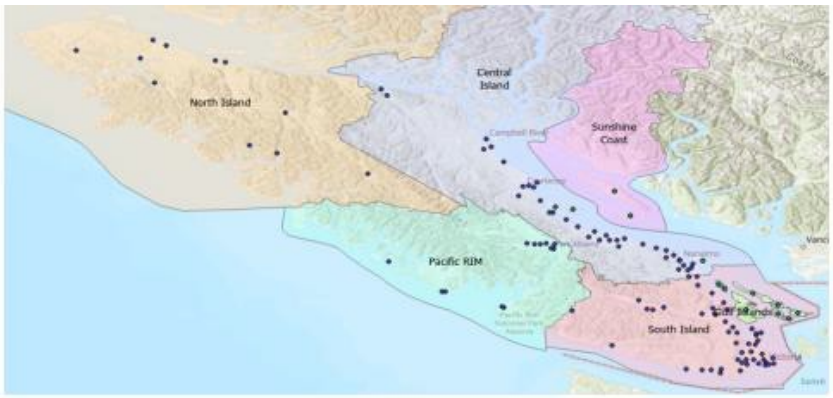


Figure 3. Communities and regional division of Vancouver Island used in the RCRSD model (Chang et al., 2020))

Framework for Altering Combinations of the Input Variables

Figure 4 illustrates the framework used to vary the models for the sensitivity analyses. Since there is an infinite number of possible infrastructure damage combinations arising from a Cascadia earthquake, the (CI) model (Bell & Bristow, 2020) was used to generate two scenarios: Case A of *Partial disconnection*, and Case B with *Extensive disconnection*. The sensitivity analyses were conducted primarily on Case A in this study, with a few runs run as an extension to examine how the more severe Case B might alter some key outcomes.

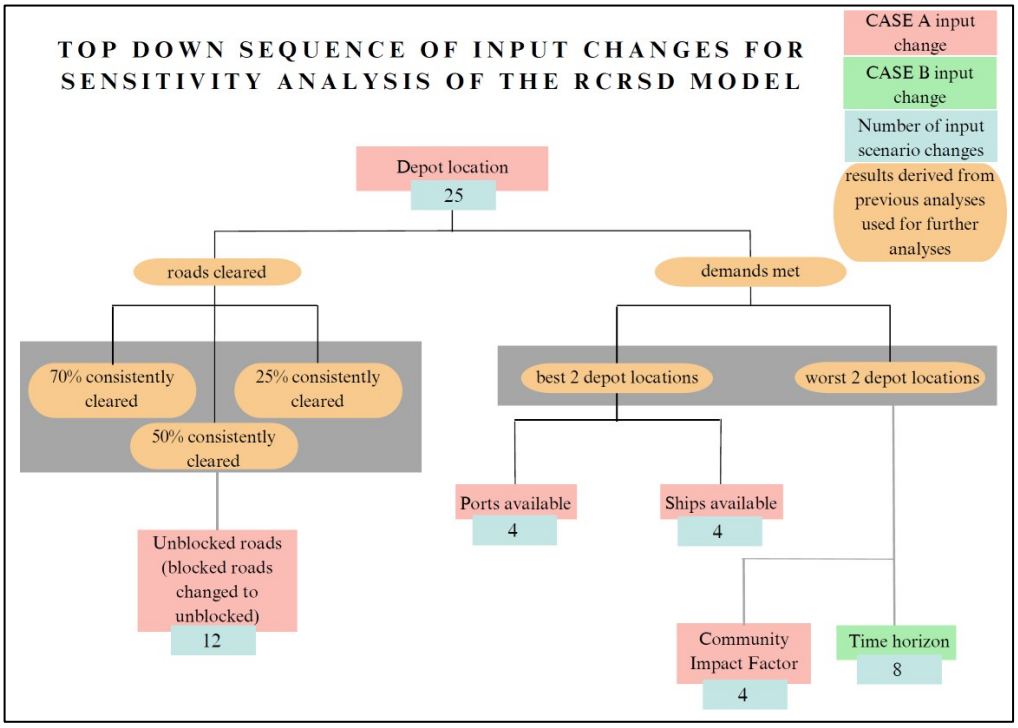


Figure 4. Sequence and number of input changes done for the sensitivity analyses

The depot location where the road clearing teams are co-located can have a significant effect on the efficiency of getting damaged roads reopened thus connecting more communities within the modelled time horizon, hence this variable is extensively studied in the sensitivity analysis, comparing 25 different base locations as noted in Figure 4. A previous study has shown that four road-clearing teams at one station is a reasonable limit, as there are greatly diminishing returns in having more (Souza Almeida et al., 2022b). One modelling outcome of particular interest is which road segments are most critical in quickly reestablishing connectivity between communities. Since the optimized set of cleared roads varies from run-to-run, one way to measure these priority roads is to observe which segments are consistently repaired across all of the computer runs. Since the concept of "consistently repaired" is subjective, three levels of consistency are defined, namely roads that are consistently cleared in at least 70% of the model runs, or 50% of the runs, or 25% of the runs respectively. Then, given each of these prescribed levels

of road repair, the models were then rerun assuming that those consistently repaired roads were not damaged to see what would be the next priority segments for repair (Figure 4).

Concurrently, the RCRSD runs were used to identify the two best depot locations, as well as the worst two depot locations (Figure 4, middle right). Even though only one depot is considered in practice in this model, comparing the best two depot locations allows the decision makers to consider choosing either of them based on other practical reasons, especially if their effectiveness is similar. Considering the worst two depot locations provides insights on how important depot location is on VI for effective goods supply (i.e. comparing the best vs the worst depot location), while also highlighting the deleterious impact of making a poor decision about this risk mitigation strategy.

For each of the best two depot locations, sensitivity was conducted on the base Case A conditions of which ports on the island are open or closed throughout the planning horizon, and which ships routes are operational or not during that period due to navigational constraints (Figure 4, bottom right). Finally, for the best and worst two depot locations respectively, the outcomes for the scenario with *Extensive disconnection* (Case B) are contrasted with those for *Partial Disconnection* (Case A). For case B, extensions in the 3-day time horizon are also examined, since very few roads are repaired within the 72-hour response window given the severe Case B conditions.

INPUT DATA

Inputs Related to Roads

The RCRSD model has a total of 2543 nodes and each node corresponds to a geographic location on or around Vancouver Island as shown in Table 2 (Souza Almeida et al., 2023).

Table 2. Types and locations of nodes used in the RCRSD model

Node range	Node type	Location
0 – 108	Community	VI
109 – 2414	Road intersection	VI
2415 – 2421	Ferry port	VI
2422 – 2428	Ferry/barge port	VI
2429 – 2457	Barge dock	VI
2458 – 2470	Community	Small islands
2471 – 2477	Ferry port	Mainland Vancouver
2478 – 2480	Ferry/barge port	Mainland Vancouver
2481 – 2515	Barge port	Mainland Vancouver
2516 – 2536	Airport/Heliport	Mainland Vancouver
2537 – 2542	Port community	VI

Port communities are the communities on VI that have a port nearby from which they can receive supplies directly via ship or barge (i.e. no road clearing required). Out of the 3707 road segments used in the RCRSD model, 1123 are assumed to be blocked in this study, with the available roads and the blocked roads for Case A shown in Figure 5. Assuming that the unblocking rate is constant for all the road-clearing teams, the unblocking time is calculated for each damaged road segment based on its length.

Inputs Related to Shipping

For the Case A Cascadia earthquake scenarios, the RCRSD model considers that there are 57 available ships, with 35 ferries and 22 barges. Each ship has its own capacity of truck containers with emergency supplies that it can carry from mainland Vancouver to Vancouver Island. Based on the structural damage model, each port is associated with an opening time, from 0 hours which indicates that it was not closed due to the earthquake, to 72 hours, which means that it is not open at all during the emergency response phase (i.e. the 3-day time horizon). For Case A, the ferry ports that are available are shown in Figure 6, while Figure 7 represents the barge docks. The barge docking locations are always available because it is assumed that barges do not require any particular infrastructure to unload supplies.

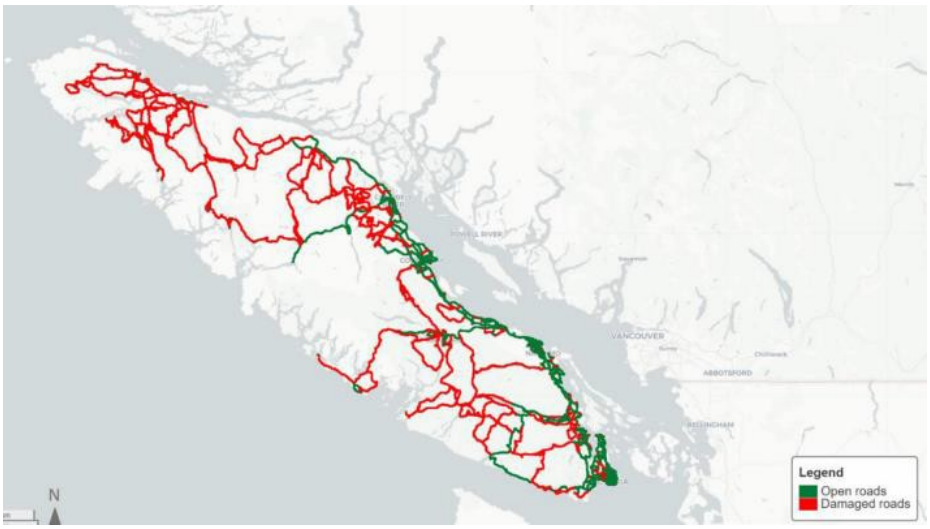


Figure 5. Open and blocked roads for Case A (source: Souza Almeida et al., 2023)

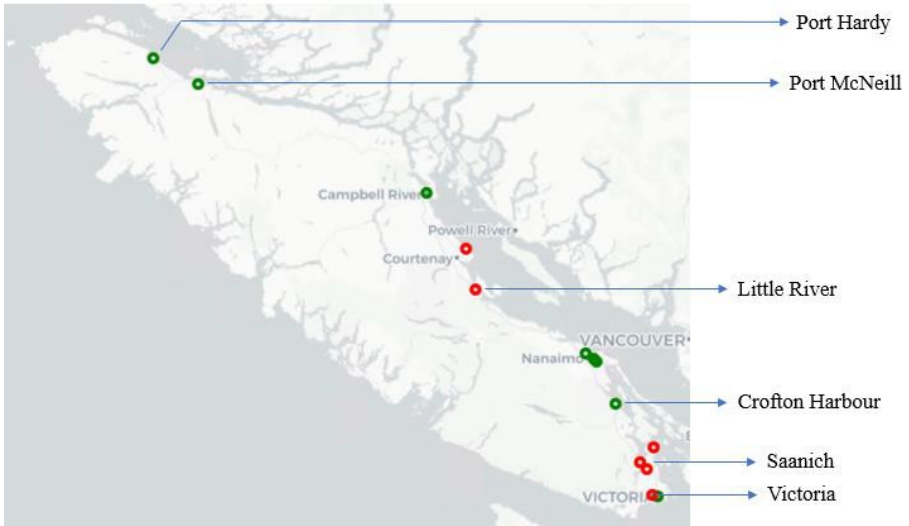


Figure 6. All ferry ports for Case A (red ones damaged)

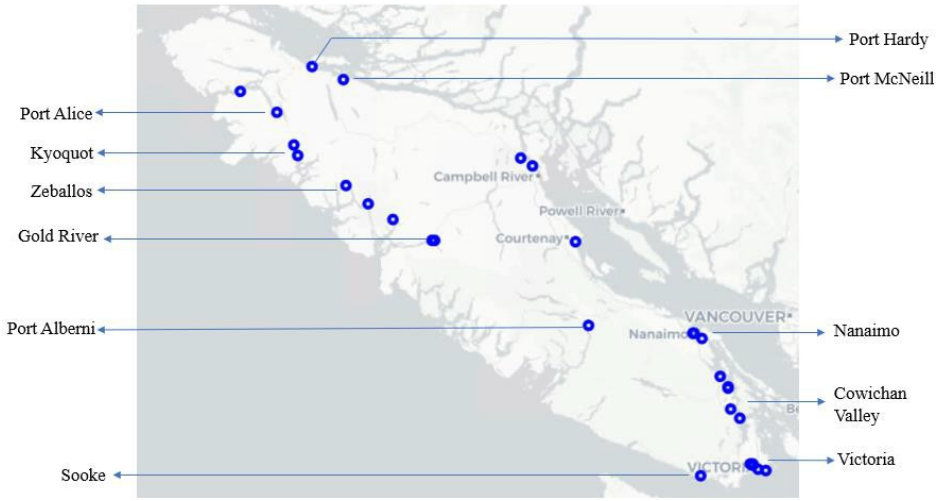


Figure 7. All barge docks for Case A

Inputs Related to Community Resilience

The resilience of communities is captured by assigning their respective nodes a value between 1 and 5. For the analysis that is done without accounting for the CIF (Community Impact Factor), all these values are changed to 5, thus giving the same priority for all communities. The CIF is used as a weight for each community in the optimization model, where the objective function is a combination of community population and CIF.

Case B Inputs

The structure of inputs for Case B is same as that of Case A, but given this more severe CI damage scenario, the roads that are considered blocked in Case B are shown in Figure 8 (Souza Almeida et al., 2023).

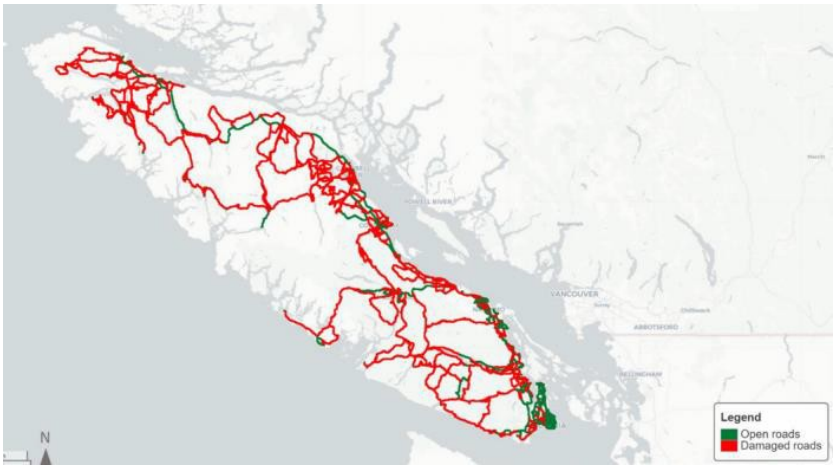


Figure 8. Open and blocked roads for Case B

RESULTS

Depot Changes Results

The first step in the sensitivity analysis for Case A was to change the location of the road clearing teams' depot. The different depot locations for which the RCRSD model was run are shown in Figure 9. For each scenario of input variation, the model was run three times to account for the changes in model results arising from the use of the GRASP heuristics method. The results of community demands supplied under every scenario (averaged between the three runs of same scenario) were tabulated as shown in Table 3. The total demand for each run is to deliver 777,646 boxes, equal to the total population on VI and the Gulf Islands.

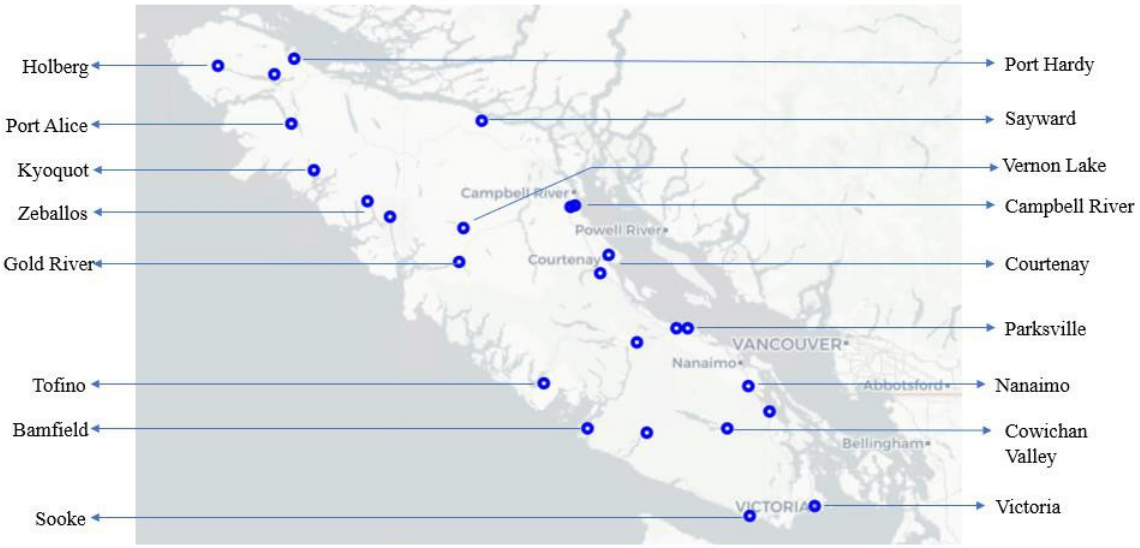


Figure 9. Alternative road clearing teams' depot locations for Case A

Table 3. Road clearing and demands supplied observations when depot location is changed (for Case A)

Scenario ID	Depot node ID	Depot location region	% Total demand supplied	% Supplied Communities
D1	42	Central Island	57.1	54.69
D2	46	North Island	63.8	60.68
D3	57	Central Island	49.9	51.04
D4	80	North Island	43.1	33.85
D5	219	Pacific Rim	42.9	33.85
D6	220	South Island	65.7	60.16
D7	742	South Island	62.6	58.85
D8	778	Central Island	61.1	57.55
D9	802	Central Island	47.8	48.70
D10	1075	Central Island	47.8	49.74
D11	1195	Central Island	65.6	61.98
D12	1544	South Island	41.6	35.16
D13	1579	Pacific Rim	63.6	58.07
D14	1926	Pacific Rim	43.5	35.42
D15	2002	North Island	41.4	35.71
D16	2012	North Island	42.3	34.90
D17	2069	North Island	43.5	36.46
D18	2165	North Island	43.0	34.64
D19	2227	North Island	43.0	35.16
D20	2313	Pacific Rim	43.3	35.94
D21	2324	Pacific Rim	44.5	38.28
D22	2343	South Island	31.6	24.74
D23	2375	South Island	63.8	60.16
D24	2395	South Island	59.3	53.65
D25	2431	Pacific Rim	53.4	51.82

From the results shown in Table 3, the best depot locations (Scenario IDs: D6, D11) and worst depot locations (Scenario IDs: D22, D15) were mapped (see Figure 10). The ranking is done based on the demands supplied and not the number of roads cleared because for certain depot locations, although there is a high degree of road clearing activity, the demands supplied and the reach to communities is very low. The demands met by the different modes of transport for the best depot location (Scenario ID: D6) and the worst depot location (Scenario ID: D22) are also shown in Figure 11 and Figure 12 respectively.



Figure 10. Best and worst depot locations

```
Total_truck_demand = 499672.0
Total_heli_demand = 8455.0
Total_ferries_demand = 1076.0
Total_barges_demand = 12342.0
Total_demand = 521545.0
Total_demand_without_heli = 513090.0
```

Figure 11. Demands met for D6 by different modes

```
Total_truck_demand = 198313.0
Total_heli_demand = 8518.0
Total_ferries_demand = 1076.0
Total_barges_demand = 1274.0
Total_demand = 209181.0
Total_demand_without_heli = 200663.0
```

Figure 12. Demands met for D22 by different modes

The results by regional division of VI are shown in Table 4, where it is seen that having the road clearing teams' depot located at either Central Island or South Island yields better results compared to having it located at Pacific Rim or North Island.

Table 4. Region-wise depot location details in terms of total demand met

Depot location region	Scenario IDs used for aggregation	Mean % demand met by region
Central Island	D1 D3 D8 D9 D10 D11	54.88
North Island	D2 D4 D15 D16 D17 D18 D19	45.74
South Island	D6 D7 D12 D22 D23 D24	54.08
Pacific Rim	D5 D13 D14 D20 D21 D25	48.53

At a higher level of aggregation, the demand results averaged for each region of VI (and the Island overall) are shown in a bar chart in Figure 13.

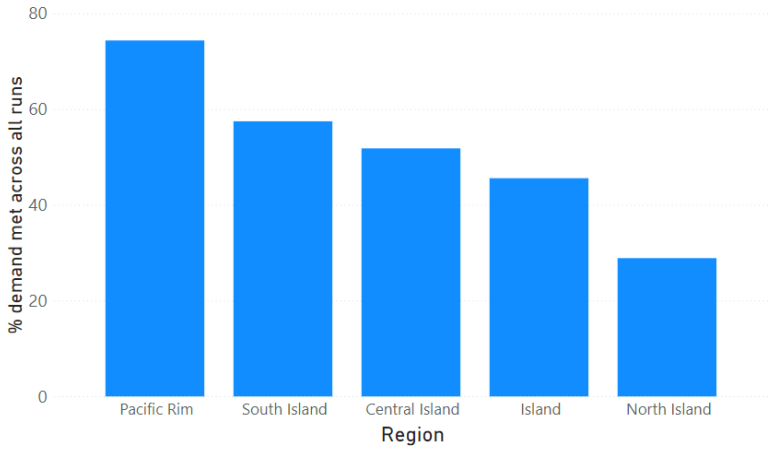


Figure 13. Regional demands satisfied across runs D1 to D25

Across the 25 input scenario variations, the communities on VI that received over 95% of their demands satisfied are shown in Figure 14. The communities that received under 5% of their demands met are shown in Figure 15. The communities in the Pacific Rim region of VI had most of their demands satisfied. The communities in the North Island region were the least likely to receive supplies.

Community	Region	Total_demand	Average_demand_met_across_all_runs_in_%
65	Pacific Rim	168	100.00
66	Pacific Rim	168	100.00
84	South Island	301	100.00
107	North Island	253	100.00
11	South Island	815	97.46
12	South Island	815	97.40
13	South Island	1113	96.00
3	Pacific Rim	1935	95.84
14	South Island	1113	95.84

Figure 14. Communities that received over 95% of their demands

Community	Region	Total_demand	Average_demand_met_across_all_runs_in_%
54	Central Island	1678	0.00
44	South Island	38050	0.32
43	South Island	38050	0.32
45	South Island	38050	0.39
41	North Island	2337	1.08
82	North Island	2409	2.07
40	North Island	664	2.60
83	North Island	2409	3.31

Figure 15. Communities that received under 5% of their demands

Roads Cleared and Subsequent Analyses

For the runs done by changing the depot locations of the road clearing teams as shown in Table 3, it was found that out of the 25 input scenarios there were no roads cleared for 5 scenarios (Scenario IDs: D4, D5, D12, D14, and D16). The reason why no roads were cleared was because the depot locations were too distant from any port or community to be reconnected within 72 hours. For example, the depot location for scenario D5 is shown in Figure 16, where the depot location is surrounded by only damaged roads, from which there is no completion of any connecting road clearing activity within the time window.

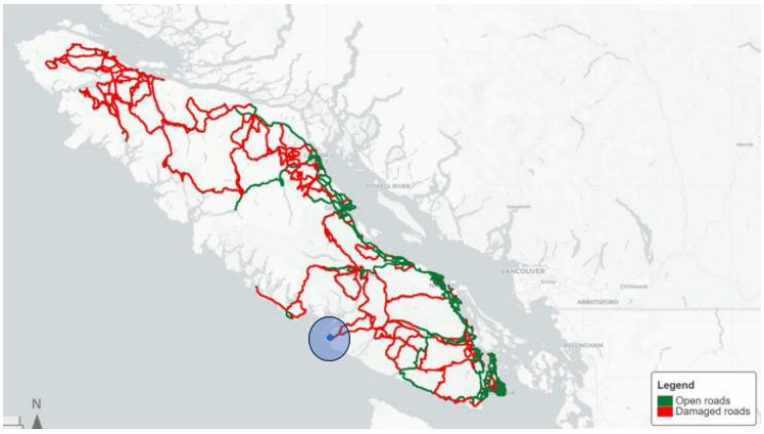


Figure 16. Map showing depot location of scenario D5 amidst blocked roads

The road segments that were cleared for each run were separately counted and for the runs with at least some road clearing activity, an average of 49 blocked road segments were cleared. The highest cleared was 81 road segments (out of 1123 damaged), associated with depot location at node 220. From all the runs, the outputs are then analyzed to find the road segments or edges that are frequently unblocked.

When all the runs are compared, it is seen that some of the blocked roads are cleared more frequently across the input scenario variations shown in Table 3 than other road segments. Figure 17 shows the road segments that were cleared for at least 70% of the original runs. This set of roads is referred to as RC1 in this study. The road segments that were most consistently cleared all lie in the South Island region. Similar results were generated for the situation with roads that were cleared for at least 50% of all the runs (referred to as RC2, see Figure 18), and for those cleared for at least 25% of all the runs (referred to as RC3, see Figure 19). For the last case, the map shows that there are significant road clearing activities happening in South Island and a few roads being fixed consistently in Central Island.



Figure 17. Roads cleared for at least 70% of the runs (RC1)

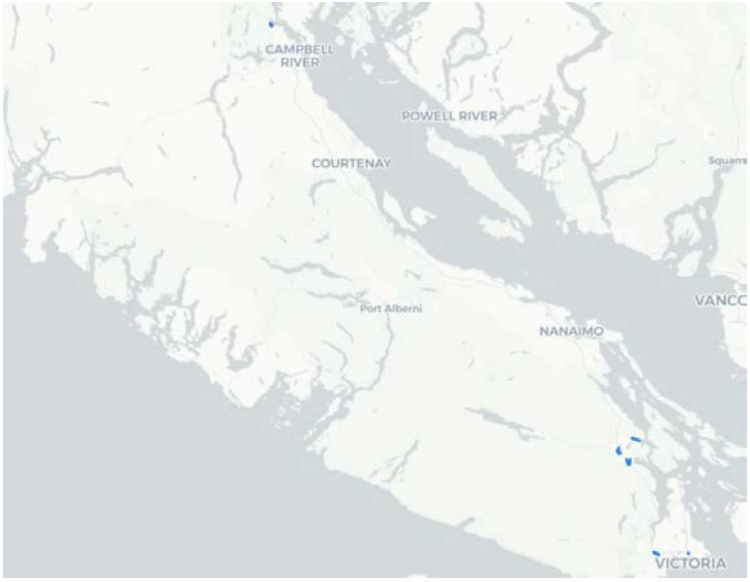


Figure 18. Roads cleared for at least 50% of the runs (RC2)

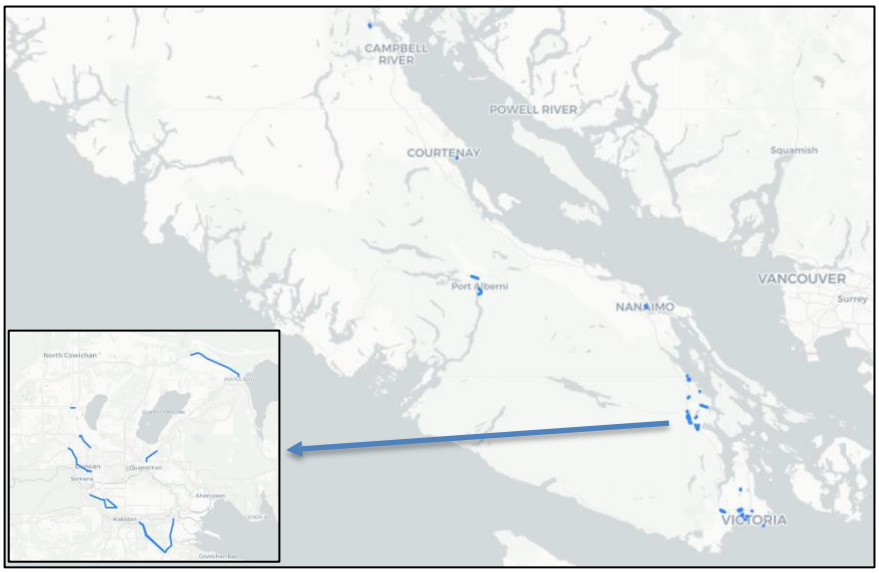


Figure 19. Roads cleared for at least 25% of the runs (RC3)

The consistently cleared roads shown in these three figures were then used as input changes for further analyses of the RCRSD model. For the 2 best depots (D6 and D11) and the two worst depots (D15 and D22), the blocked roads from the original dataset that are consistently cleared were considered to be unblocked as shown in Table 5 (at RC1, RC2 and RC3 levels respectively).

Table 5. Input variable scenario changes for consistently unblocked roads

	Input Variable chosen for analysis	
Run ID	**Depot location**	**Consistently unblocked roads**
D6RC1	220	RC1
D11RC1	1195	RC1
D15RC1	2002	RC1
D22RC1	2343	RC1
D6RC2	220	RC2
D11RC2	1195	RC2
D15RC2	2002	RC2
D22RC2	2343	RC2
D6RC3	220	RC3
D11RC3	1195	RC3
D15RC3	2002	RC3
D22RC3	2343	RC3

Table 6 shows the number of boxes that are supplied for each scenario, as well as the percentage of the total demand that is met, as the average of 3 runs done for each scenario. The total demand for all communities for each run is 777,646 boxes of supplies (i.e. the population of Vancouver Island and the Gulf Islands). The results can be contrasted with the original results reproduced in the final two columns. For the depots in the best locations (D6 and D11), there are no significant changes in the demands satisfied when the roads represented by RC1, RC2, and RC3 are considered undamaged. However, for the depot's worst locations, D15 and D22, that earlier showed demands satisfied as less than or around 40%, when the consistently cleared ones are considered unblocked, there are significant improvements in the demands satisfied as well as the number of communities that receive supplies.

Table 6. Results observed for changes in unblocked roads

				Original results	
Scenario ID	**Demand met in number of boxes**	**% Total demand met**	**% Communities supplied**	**Scenario ID**	**% Total demand met**
D6RC1	491645	63.22	61.33	D6	65.7
D11RC1	485185	62.39	58.98	D11	65.6
D15RC1	486428	62.55	57.42	D15	41.4
D22RC1	438715	56.42	52.01	D22	31.6
D6RC2	453078	58.26	55.08	D6	65.7
D11RC2	504896	64.93	60.94	D11	65.6
D15RC2	477331	61.38	56.64	D15	41.4
D22RC2	484570	62.31	58.98	D22	31.6
D6RC3	526870	67.75	60.55	D6	65.7
D11RC3	522386	67.18	62.89	D11	65.6
D15RC3	507548	65.27	60.94	D15	41.4
D22RC3	525983	67.64	62.89	D22	31.6

Critical Ports and Subsequent Analysis

The location of the most critical barge dock based on box supplies' origin is shown in Figure 20. For the communities serviced by trucks, 55% of the supplies went from this location. In comparison, the busiest ferry port, shown in Figure 21, contributed to only about 6% of the total supplies dispatched from the ports.

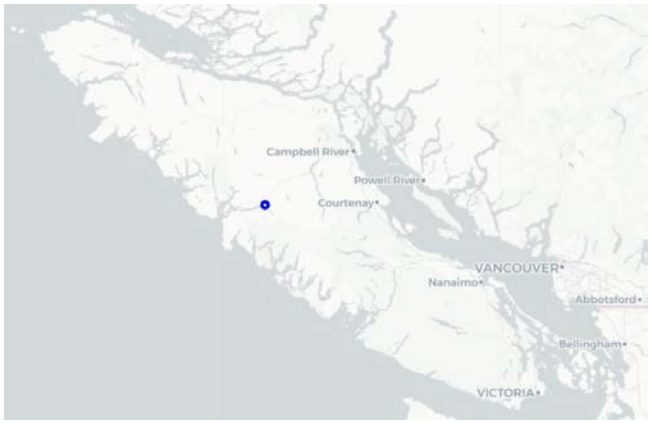


Figure 20. Location of critical barge dock

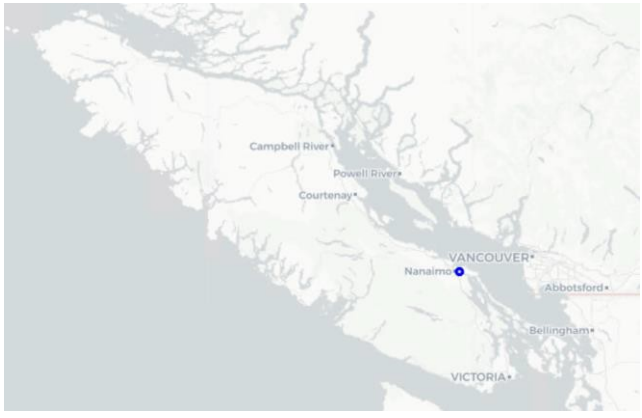


Figure 21. Location of critical ferry port

The next set of analyses were done by considering having either of the critical ports closed. The various input scenarios for which this analysis was done is shown in Table 7. For the two depots with the best locations (scenarios D6 and D11), combined with either the barge port closed (P1) or the ferry port closed (P2), the impact on deliveries are shown in Table 7, with the original results shown in the last two columns. For the scenarios where the barge port is considered closed (D6P1 and D11P1), the total demands satisfied are significantly less when compared to the scenarios where they are considered open (scenarios D6 and D11). The last two rows show that the impact of having the ferry port closed is much less, as it delivers a small proportion of the supply in the first place.

Table 7. Results observed for changes in ports availabilities

				Original results	
Scenario ID	**Demand met in number of boxes**	**% Total demand met**	**% Communities supplied**	**Scenario ID**	**% Total demand met**
D6P1	97075	12.48	21.61	D6	65.7
D11P1	26435	3.40	13.80	D11	65.6
D6P2	453662	58.34	53.91	D6	65.7
D11P2	464374	59.72	55.73	D11	65.6

Availability of Ships

Since the barge deliveries are greater than ferry deliveries, the RCRSD model is run by considering only 50% of the barges to be available, considering the two best depot locations. The case where all ferries are available and only 50% of barges are available is labelled as S1 in Table 8, while the case with all ferries unavailable and all barges operational is S2. When the barges are reduced by 50%, the total demand satisfied is reduced by a significant amount (comparing for example the original D6 result on the right, with the new D6S1), where the demand satisfied drops from 65.7% to only 12.1%. In contrast, for scenarios D6S2 and D11S2 where all ferries are assumed non-operational, the demands met are reduced by only about 10% from the runs D6 and D11 where all ferries are in use.

Table 8. Results observed for changes in ships' availabilities

				Original results	
Scenario ID	**Demand met in number of boxes**	**% Total demand met**	**% Communities supplied**	**Scenario ID**	**% Total demand met**
D6S1	94104	12.10	14.50	D6	65.7
D11S1	147291	18.90	21.58	D11	65.6
D6S2	435185	55.96	51.56	D6	65.7
D11S2	437010	56.20	51.82	D11	65.6

Community Impact Factor (CIF) Change

For the two best (D6 and D11) and two worst (D15 and D22) depot locations, new scenarios are run where the CIFs are neglected by making them all the same value, represented by C1 in Table 9. Comparing the results with the original outputs given in the two rightmost columns, there are no significant changes in the total demands met (greatest variation is 3%), except scenario D11C1 where total demand served dropped from about 65% to 47.68%. However, there is a decrease in the percentage of communities that receive supplies when the CIF is neglected in all scenarios.

Table 9. Results observed for changes in CIF

					Original results	
Scenario ID	**Demand met in number of boxes**	**% Total demand met**	**% Communities supplied**	**Scenario ID**	**Scenario ID**	**% Communities supplied**
D6C1	483812	62.22	53.51	D6	65.7	60.16
D11C1	370800	47.68	42.44	D11	65.6	61.98
D15C1	327414	42.10	34.47	D15	41.4	35.71
D22C1	261209	33.59	24.08	D22	31.6	24.74

Case B Findings

For the extensive disruption scenario (Case B, see Figure 4), the RCRSD model is run for only the two best and two worst depot locations from Case A analysis. All of the B scenario runs resulted in the same number of boxes being delivered. Changing the depot location does not make a difference in the results because there is no clearing activity completed in any in the B scenarios, so no communities are being reconnected to ports or barge docks within 72 hours. Out of the supplies that were dispatched, 5% was done by helicopters, 10% by barges directly, less than 1% by ferries directly, and about 84% by trucks of which 79% originated from the port in Victoria. When the time horizon was increased to 360 hours (15 days), although some road clearing happens, there is only about a 10% increase in the demands supplied when compared to the time horizon of 72 hours. Further details are provided in Vattoni (2023).

DISCUSSION

These sensitivity analyses on the distribution of relief supplies on Vancouver Island have explored the impacts of changes in some of the key factors affecting the outcome. While detailed outputs for each of the model runs are conserved, namely which roads are repaired and which communities receive supplies, better insights are provided through the use of regional aggregation of results, maps, and tables showing percentages of achievement. In response to the questions posed in the introduction section, select observations from these runs are presented in the following paragraph, while more extensive sensitivity analyses can be conducted eventually on a broader range of scenarios with varying degrees of earthquake damage.

The partial disconnection scenario (Case A) showed that the best location to have the depot would be the Central or the South Island regions of VI, but the latter is more likely to suffer damage from the earthquake. It was also shown that if the roads on South Island were reinforced so that they stayed unblocked, supplies are able to reach a larger percentage of the population. On changing the inputs related to shipping such as the ports that are open or the number of ships available, it was seen that over 50% of the supplies to communities went from a single barge dock halfway up the west coast. Additionally, when the critical roads were considered undamaged, it was found that the supplies to communities increased because the roads to this barge dock were cleared. These findings would help decision makers ensure to the degree possible the availability and strengthening of this infrastructure. Eliminating the Community Impact Factor (CIF) from the calculations does not affect the total supply significantly, but the number of communities that receive supplies are fewer. Preparedness experts can use these findings to help advocate for improved resilience (strengthening community infrastructure, maintaining more emergency supplies), particularly for certain towns which consistently receive less than 2% of their demand across all model runs (with those results housed in the detailed output files). The main findings from the more extreme Case B analysis is that 3 days is insufficient for conducting any appreciable road clearing, thus longer waits for deliveries are to be expected, or additional air response would be required.

ACKNOWLEDGMENTS

We gratefully acknowledge the helpful comments from the reviewers. This study was financially supported by Marine Environmental Observation Prediction and Response (MEOPAR) Network of Centres of Excellence of Canada, and from Emergency Management British Columbia (EMBC).

REFERENCES

Bell, A. M., & Bristow, D. N. (2020). Modelling the marine transport resilience of Vancouver Island in a Cascadia subduction zone earthquake scenario. *Sustainable and Resilient Infrastructure*, *7(5)*, 1–18. https://doi.org/10.1080/23789689.2020.1845482

Chang, S., Bristow, D., Goerlandt, F., Pelot, R., Goodchild, A., Lin, C., & Zhou, L. (2020). *Planning for a catastrophic earthquake in British Columbia: Marine transportation disruption and coastal community resilience*. SIREN Project Report, p. 112. https://www.maritimeriskandsafety.ca/

Chang, S. E., & Tanner, A. (2022). A community impact scale for regional disaster planning with transportation disruption. *Natural Hazards Review, 23(3)*, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000563

Clague, J. J. (2002). The earthquake threat in southwestern British Columbia: A geologic perspective. *Natural Hazards, 26*, 7–34.

Natural Resources Canada. (2010). Earthquakes in SW British Columbia. Earthquakes Canada (nrcan.gc.ca)

Rogers, C. G. (1998). Earthquakes and earthquake hazard in the Vancouver area. Geology and Natural Hazards of the Fraser River Delta, British Columbia. In J.J. Clague, J.L. Luternauer, & D.C. Mosher (Eds.); *Geological Survey of Canada*, Bulletin 525, 17-25.

Souza Almeida, L., Goerlandt, F., & Pelot, R. (2022a). Trends and gaps in the literature of road network repair and restoration in the context of disaster response operations. *Socio-Economic Planning Sciences, 84*. https://doi.org/10.1016/j.seps.2022.101398

Souza Almeida, L., Goerlandt, F., Pelot, R., & Sörensen, K. (2022b). A greedy randomized adaptive search procedure (GRASP) for the multi-vehicle prize collecting arc routing for connectivity problem. *Computers & Operations Research, 143.* https://doi.org/10.1016/j.cor.2022.105804

Souza Almeida, L., Rodrigues, L., Goerlandt, F., Ancona-Segovia, J., Pelot, R., Bristow, D., & Chang, S. (2023). Datasets of disrupted transportation networks on Canada's West Coast in a plausible M9.0 Cascadia Subduction Zone earthquake scenario. *Data in Brief, 46*. https://doi.org/10.1016/j.dib.2022.108762

Vattoni, P. (2023). Sensitivity analysis of a road clearing and relief supplies distribution model for Vancouver Island in the event of a Cascadia megathrust earthquake, MASc thesis, Dalhousie University. https://dalspace.library.dal.ca/handle/10222/82919



Open Track

Track Chairs

Michael Middelhoff, University of Münster

Kees Boersma, Vrije Universiteit Amsterdam

R2DN – Integrating Databases about Natural Disasters

José Veríssimo
PPGI - Graduate Program in Informatics
Universidade Federal do Rio de Janeiro
verissimo@ufrj.br

Marcos R. S. Borges
PPGI - Graduate Program in Informatics
Universidade Federal do Rio de Janeiro
mborges@ppgi.ufrj.br

ABSTRACT

The historical data on disasters serves various purposes, such as managing the risk of future disasters, assessing risk in the insurance sector, and formulating policies to reduce disaster risks, among others. To meet these objectives, there are databases, covering several types of disasters. However, in addition to differences in purposes, scope, access, and coverage, there are variations in the criteria for registering a disaster in the database and the information used for this register. Due to these differences, we have a scenario with discrepancies of information across existing databases. In this study, we propose a repository that combines the main advantages of some global databases, the data needs due to multilateral agreement, and the possibility of integrating regional or specific databases. The proposed Natural Disaster Data Repository (R2DN) contains a data schema and a data mapping process with two of the main global databases (EM-DAT and DesInventar) and S2ID.

Keywords

Disaster databases, database integration, EM-DAT, DesInventar, S2ID

INTRODUCTION

Natural disasters are phenomena that have been part of the Earth's dynamics for centuries, with recorded human casualties and material damages. When they occur in uninhabited areas, they are termed natural events. Earthquakes, floods, tsunamis, epidemics, and droughts are some of these phenomena with distinct characteristics and consequences. In some cases, an event may be associated with one or more of these phenomena, such as the earthquake that triggered the Indian Ocean tsunami in December 2004, affecting 13 countries in the region.

The Centre for Research on the Epidemiology of Disasters – (CRED, 2020b), at the Université Catholique de Louvain in Belgium, estimated that, in the last two decades (2000 to 2019), 7,348 disaster events (excluding biological and technological disasters) were recorded, resulting in approximately 1.23 million lives lost, affecting over 4 billion people, and causing economic losses of around $2.97 trillion. These figures represent a significant increase compared to the two previous decades (1980 to 1999), which recorded 4,212 events, 1.19 million deaths, over 3.2 billion people affected, and losses of approximately $1.63 trillion, as summarized in **Figure *1***.

In a scenario where various local databases provide information for the purposes of their regions, sharing information between databases is necessary for more comprehensive studies that include multiple regions. To ensure the agility, quality, and integrity of data, compatibility between the data is necessary, allowing for sharing between databases. However, as mentioned by (Panwar & Sen, 2020), existing major databases still exhibit disparities in disaster recording, making data exchange for comparison and analysis challenging. According to (Koç & Thieken, 2018), the lack of reliable, consistent, and comparable data is seen as a significant obstacle to effective long-term prevention of losses.

The Emergency Events Database (EM-DAT), developed by CRED in 1988, is widely recognized as a reference for the academic community, industry, and government entities in the disaster assessment phase. EM-DAT is maintained by the Université Catholique de Louvain with the support of the World Health Organization (WHO). It is available to non-commercial individuals and organizations upon registration on its website (CRED). A simple search in the Scopus database reveals that over a hundred articles, since 2012, reference EM-DAT in their abstracts. There are other databases, public and private, enabling comparisons between their characteristics.

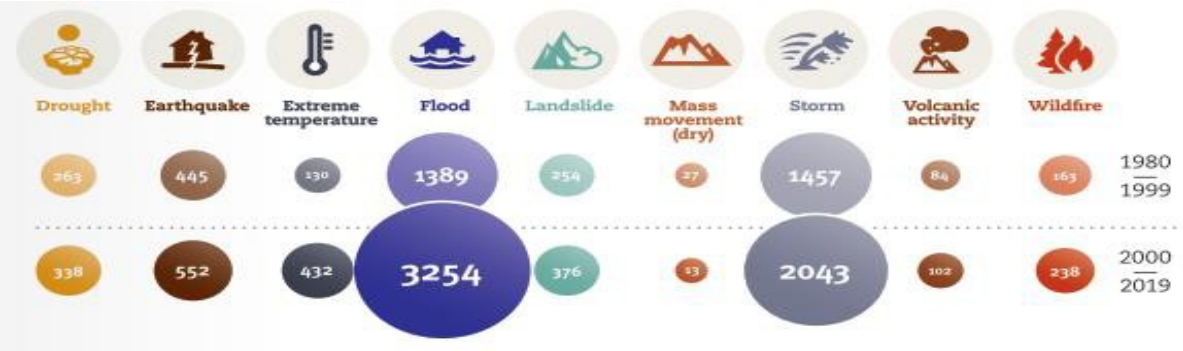


Figure 1. Impact of Disasters: 1980-1999 versus. 2000-2019 (CRED, 2020b)

The DesInventar, an initiative of UNDRR (United Nations Office for Disaster Risk Reduction), is a system designed for the acquisition, recovery, querying, and analysis of disaster records. This system is made available to countries in Latin America, the Caribbean, Asia, and Africa.

The S2ID is an Integrated Disaster Information System developed by CEPED/UFSC in 2011/2012 ("Ministério do Desenvolvimento Regional - S2ID"), serving as a support tool for states and municipalities affected by disasters that require federal assistance. The system enables these federative entities to exchange information with the federal government to obtain necessary resources to assist the affected population.

This work supports the proposal that each country or institution should have its own specific database of natural disasters to meet the specific needs of understanding natural phenomena occurring in its territory and their impacts on society. We share the proposals presented by (Wirtz, Kron, Löw, & Steuer, 2014) and (Khankeh, 2022). The purposes of creating and using disaster damage databases, according to (Khankeh, 2022), should include:

1. Promoting disaster relief, recovery, and reconstruction programs.
2. Assessing risks of future disasters.
3. Estimating the economic feasibility of investments made to reduce losses.
4. Monitoring disasters and their impacts to achieve international goals established for disaster risk reduction.
5. Conducting thematic analyses, such as differences in deaths by gender or age group in disasters.

However, we understand that the proliferation of specific databases should consider adhering to some standards already established by other institutions managing disaster databases. This approach can facilitate data sharing among them and enable the exchange and/or crossing of temporal and geographical information necessary for generating more effective analyses supporting governments and organizations in decision-making and the formulation of disaster risk reduction policies. The exchange and sharing of disaster information among countries are explicitly stated in Priority 4 of the Sendai Framework for Disaster Risk Reduction 2015-2030 (UN, 2015), and it is part of the commitment made by 187 UN member countries.

It is also important to emphasize the need to include data in the databases grounded on the demand for new attributes aligned with the disaster risk reduction indicators required by the Sendai Framework. Divergences between databases hinder data sharing and, consequently, limit the potential utilization of disaster databases. This is the main problem that will be addressed in this work.

What is the structure of a data repository that allows the assessment of natural disasters, is compatible with global databases, meets specific demands and complies with multilateral agreements in disaster management?

To answer this question, we established three goals that are described and discussed in this work. They are:

1. To develop a database structure that integrates disaster records from different databases.
2. To include the necessary information for the Sendai Framework indicators.
3. To enable the integration of specific databases aimed to support specific purposes from countries or institutions to assist their policies.

In the following sections, we will present the work context, the methodology employed to develop the study, the description of the three databases selected for the integration exercise, the criteria used for selection, the taxonomy and the R2DN data model, as well as an instance of the integration of the three described databases.

BACKGROUND AND METHODOLOGICAL APPROACH

The growing global demand for actions aimed at disaster risk reduction has driven the need for databases capable of storing and retrieving information related to these events. In response to this scenario, we observe a significant number of published articles addressing the use of disaster databases for various purposes, regardless of the authors' areas of expertise.

In Table 1, we present the results of a search conducted in the Scopus database. This research involved the inclusion of the keyword "database" along with terms representative of each phase of disaster management. We also provide the number of articles found since the year 2013, offering an overview of academic production in this field.

Table 1. Number of articles found from 2013 onwards, in the Scopus database, using the corresponding sentences Data retrieved on 10/31/2023

Search sentence	Number of articles
((“natural disaster” OR “natural hazard”) AND (“prevention” OR “risk reduction”) AND (“database”))	367
((“natural disaster” OR “natural hazard”) AND (“response”) AND (“database”))	347
((“natural disaster” OR “natural hazard”) AND (“assessment”) AND (“database”))	760

The Design Science Research (DSR), widely employed in Information Systems research and advocated by (Hevner, March, Park, & Ram, 2004), is the methodology we have chosen to conduct our research. Its guidelines, which include identifying a problem in society, creating an artifact capable of contributing to the solution of this problem, evaluating the contributions generated by the artifact, and disseminating the achieved results, have shown to be consistent with the planned strategies for the development of our study.

In our study, we identified a problem related to the difficulty of data sharing among existing disaster databases. Additionally, the solution to this problem must meet the requirements for emergency response in Brazil and be in compliance with international agreements in the field of disaster risk reduction. To address this issue, we designed a repository (artifact) that integrates disaster records from three existing databases. The challenges encountered during the integration process are analyzed, and based on these analyses, new requirements are formulated to enhance the repository's design, aiming to overcome the identified challenges.

In Figure 2, we present a diagram containing the stages and deliverables of our methodological approach, aligned with the guidelines of DSR.

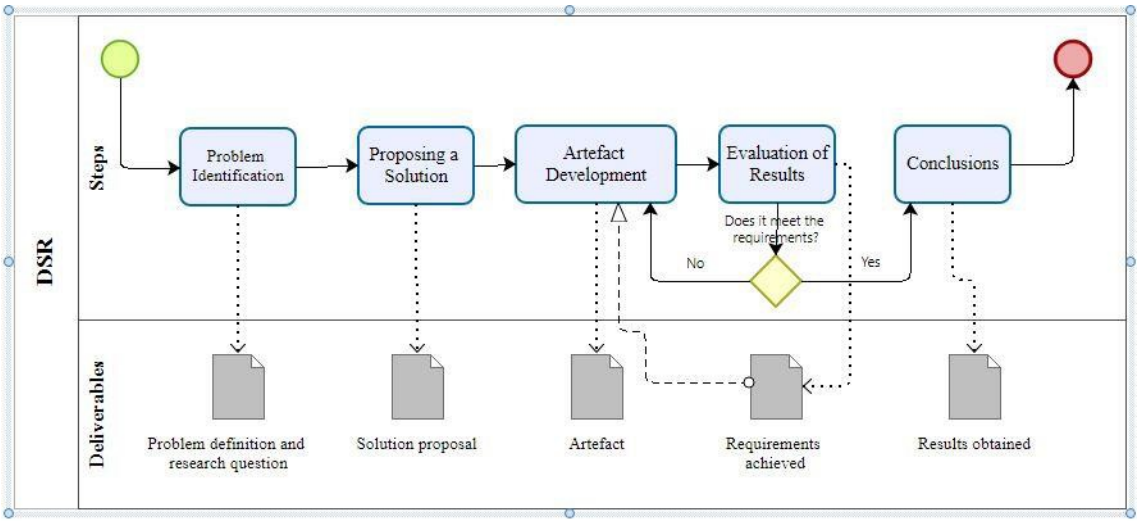


Figure 2. DSR steps and its deliverables

In the "Problem Identification" stage, where we identified a specific problem in society, it was described in this section along with the research question that contains the requirements for the desired solution. In the "Proposing a Solution" stage, we developed a model for a repository that meets the solution requirements, as presented in Section “The Proposed Model for R2DN”.

To demonstrate the feasibility of the proposed model, we implemented an instance in the "Artifact Development"

stage, as described in Section “Instantiating R2DN”. The repository development is divided into two parts. In the first part, actions necessary for the instantiation of the artifact are identified in the face of encountered difficulties, and in the second part, the implementation of R2DN is carried out, followed by an "Evaluation of the results". After each evaluation stage, the results are analyzed, and adjustments to the artifact are made as needed to achieve the expected outcomes.

DISASTER DATABASES

This work is based on three distinct disaster databases. Table 2~outlines the justification behind the selection of each one, as well as the criteria adopted for registering catastrophic events.

Table 2. Selected databases with motivation and event recording criteria

Database	Motivation	Registration Criteria
EM-DAT	Widely recognized by the disaster management community and accessible to the public.	10 or more human deaths / 100 or more affected people / Declaration of a state of emergency by the country and/or an appeal for international assistance.
S2ID	Aims to analyze emergencies in Brazilian states and municipalities.	Requesting financial assistance from the federal government.
DesInventar	Records disasters in all South American countries, except Brazil.	Each country using the system has the autonomy to establish its own criteria.

The use of different disaster databases allows for a more comprehensive understanding of disasters due to variations in criteria for recording in each database. However, these discrepancies lead to disparities between databases, making information sharing a challenge and consequently limiting the potential utilization of these disaster databases.

Despite EM-DAT being the globally recognized disaster database, we emphasize a significant discrepancy with the reality in Brazil (the country we chose to perform the data integration). The report on damages and losses resulting from natural disasters in Brazil (1995-2019), produced by CEPED (UFSC) and updated by FAPEU with support from the World Bank (FAPEU, 2020), indicates 64,429 disaster records with economic losses of R$ 333.36 billion (equivalent today to approximately US$60 billion) over these 25 years of the study. In the same period, EM-DAT recorded (query conducted on 07/20/2022) 136 disasters, totaling economic losses of approximately US$ 18 billion.

Several reasons can justify such a discrepancy, including different criteria for recording disasters and distinct purposes of the databases. While EM-DAT aims to register globally impactful disasters for humanitarian purposes, CEPED's database aims to register disasters that states and municipalities require some form of federal assistance.

The numbers presented above highlight the need for a specific disaster database for Brazil that aligns with our objectives and is compatible with major global databases. This would enable data sharing and comparative studies on the losses and damages caused by disasters across countries. We believe that recording losses from disasters is the first step in assessing resilience evolution in different regions.

According to (IRDR, 2014), most disaster databases record human and economic losses in some way, but they lack a standard set of indicators established from a common understanding. An initiative to address this issue is the set of indicators established in the Sendai Agreement.

To initiate the selection of disaster databases, we relied on the systematic review conducted by Khankeh (2022), which provided an analysis of these databases. Based on this review, we selected those with a global presence and a higher number of references on Google Scholar. Table 3 presents some of the characteristics collected from these databases' websites in June 2023. Note that not all information was completed in the websites. In addition, we included the S2ID to the table because it is the official system of the Brazilian government.

The process of selecting R2DN attributes was divided into three stages:

1. Analyzing the data structures of the three selected databases.
2. Analyzing the data needs of the Sendai Framework indicators.
3. Analyzing the data organization of the DaLA methodology.

Table 3. Comparison of the characteristics of disaster databases (the numbers on the website point to the URLs at the end of the table)

	EM-DAT	NatCatService	Sigma	DesInventar	GLIDEnumber	Dartmouth	S2ID
Site	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Administrator	CRED – Belgium	Munich Re – Germany	Swiss Re - Switzerland	LA RED with support from the UN	Asian Disaster Reduction Center (ARDC)– Japan	University of Colorado	National Secretariat for Civil Protection and Defense - Brazil
Start year	1988	1980	Not available	1994	2004	1993	1988
Objective	Support humanitarian actions and provide a basis for assessing vulnerabilities and priorities.	Developed to meet the needs of the insurance industry.	Developed to meet the needs of the insurance industry.	Support analysis of disaster trends and their impacts.	Proposal for a single, global code for identifying disasters.	Remote sensing-based measurements and mapping of the digital map of surface water changes.	The official system for Brazilian states and municipalities to request federal aid.
Disasters since	1900	From 1980 onwards for disasters that meet your criteria.	1970	Varying for each country.	Not available.	Uninformed.	2003
Available	Public upon registration on the website.	Private, but publishes periodic atlases of losses in disasters.	Private, but publishes periodic atlases of losses in disasters.	Public	Public	Public	Public
Registration criteria	> 10 deaths/ > 100 affected/ Declaration of a emergency/ Request for help	Events that result in deaths, injuries, homelessness and/or property damage.	> 20 deaths, > 50 injured, > 2,000 homeless and/or insured losses > $14M.	At the discretion of the registered information sources.	At the discretion of the accredited countries.	Changing water levels on Earth (captured by remote sensing).	Declaration of a state of emergency or state of public calamity.
Number of records *	25,596 of which 16,398 were natural disasters.	Not available	Not available	Not available	7,589 disasters with Glidenumber.	Not applicable	Not available

1 - https://emdat.be
2 - https://www.munichre.com/en/solutions/reinsurance-property-casualty/natcatservice.html
3 - https://www.sigma-explorer.com/
4 - https://www.desinventar.net/
5 - https://www.glidenumber.net/glide/public/about.jsp
6 - https://floodobservatory.colorado.edu/index.html
7 - https://s2id.mi.gov.br/
* Note: All links accessed on 06/13/2023

EM-DAT's data structure is accessible through its website (CRED). The available information covers various aspects of disasters, including identification, classification, location, and human and financial impacts. Among these attributes, five are related to disaster classification, covering nature, group, subgroup, type, and subtype of the event. Regarding the disaster's location, information about the continent and country is provided. However, there is only one attribute to describe the specific location within the country. Information related to financial impacts focuses on total reconstruction costs without providing specific details about the affected economic sectors.

DesInventar provides a document addressing the system's use, giving an overview of its data structure (UN Office for Disaster Risk Reduction (UNISDR), 2019). However, there is a lack of formalism in defining this data. The disaster classification is simplified into a single attribute called "event type" while financial impacts are detailed concerning the affected economic sectors.

S2ID metadata is accessible in the Digital Atlas of Disasters in Brazil (Ceped - Centro Universitário De Estudos E Pesquisas Sobre Desastres, n.d.). Disaster classification information resembles that of EM-DAT, as it served as the basis for S2ID disaster classification. The disaster's location is clearly defined, as municipalities where events occurred register these occurrences to request financial assistance from the federal government. As a result of this process, financial impacts, especially those related to the public sector, are thoroughly documented.

Sendai Framework

The Sendai Framework is a global initiative aimed at promoting disaster risk reduction policies worldwide. It aims to measure the progress of initiatives adopted by countries between 2020 and 2030, compared to the period from 2005 to 2015, through the establishment of seven global targets (UN, 2015). To achieve this goal, the framework seeks to prevent new disasters and reduce risks by implementing integrated and inclusive measures that aim to minimize exposure to risks and vulnerability to disasters. Additionally, it aims to enhance preparedness for disaster response and recovery to strengthen the resilience of affected communities.

These targets, to be achieved by 2030, are divided into:

a. Reduce global disaster mortality.
b. Reduce the number of affected people globally.
c. Reduce direct disaster economic loss in relation to global gross domestic product (GDP).
d. Reduce disaster damage to critical infrastructure and disruption of basic services.
e. Increase the number of countries with national and local disaster risk reduction strategies by 2020.
f. Enhance international cooperation to developing countries.
g. Increase the availability of and access to multi-hazard early warning systems and disaster risk information and assessments to people.

Targets a, b, c, and d require data from national disaster databases as evidence of progress in disaster risk reduction policies. Therefore, the attributes to be included in the R2DN data model must align with this need.

THE PROPOSED DATA MODEL FOR R2DN

In this section, we present the answer to our research question. The elements that supported the construction of the model include information sources from the metadata of the three databases, in addition to the indicators of the Sendai Framework.

Figure 3 illustrates the process of modeling the R2DN data schema. The first activity, "Identify disaster data categories" aims to organize the data into categories and subcategories. This approach, combined with the metadata, is crucial for a more precise understanding of the data and assists in defining responsibilities for collecting information about disasters.

Next, in the activity "Select R2DN attributes" we chose the data based on the attributes of the selected databases that will compose the disaster record in R2DN. In the subsequent activity, named "Allocate disaster attributes to their categories" we associate each attribute with its corresponding data category. For the data selected for R2DN, we check if there is a metadata description available in the source database. If it exists, this description is integrated into the R2DN model through the "Search metadata in the source database" activity. However, if it does not exist, in the "Define metadata" activity, metadata is generated for these data.

Finally, in the activity "Improve metadata with real examples" in the face of the absence or low quality of metadata in the source databases, we resort to real examples in the databases to enrich the corresponding metadata and improve its understanding.

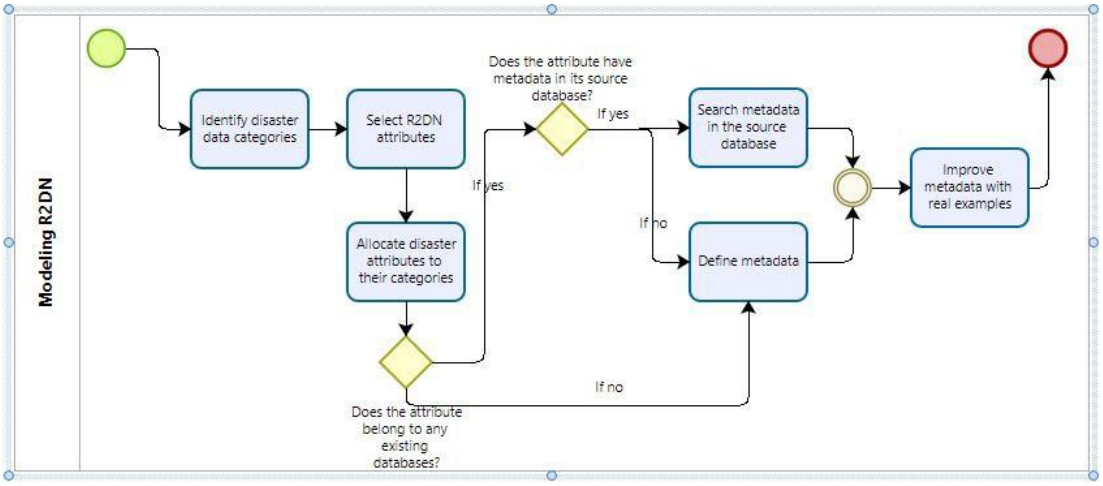


Figure 3. R2DN modeling process

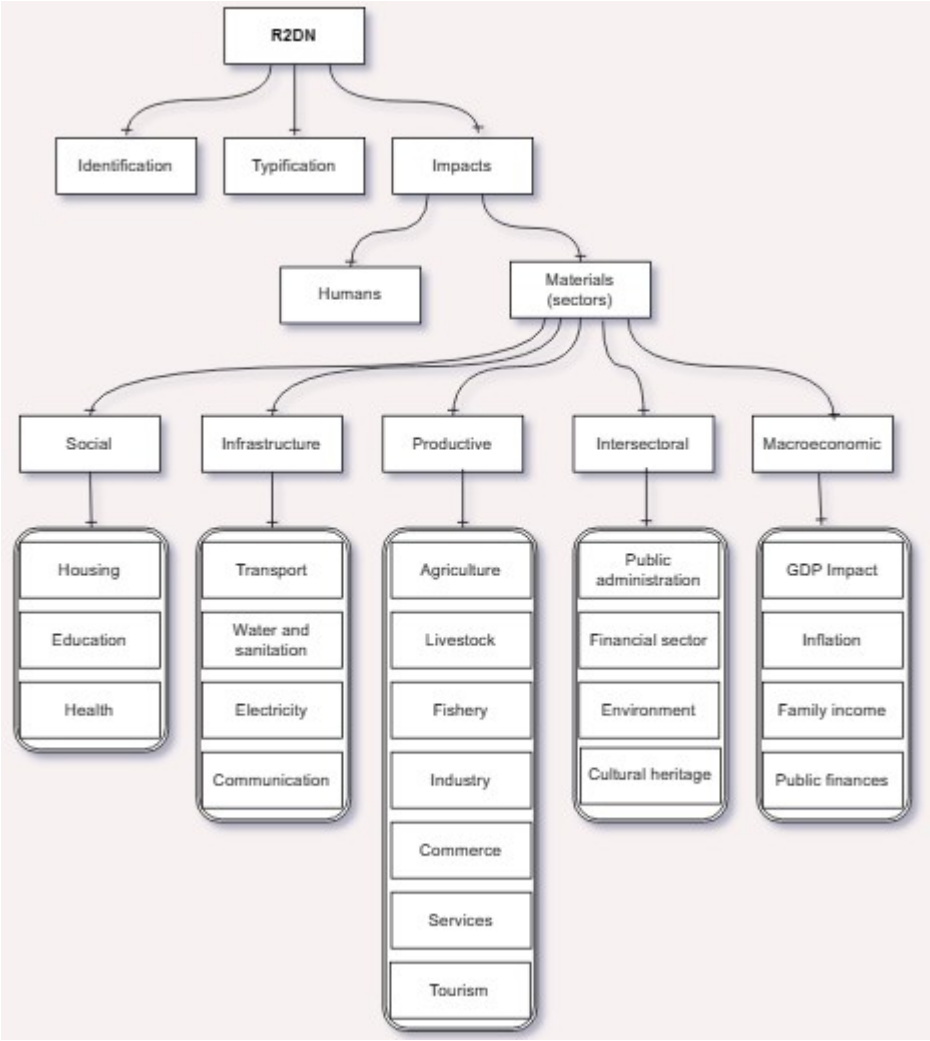


Figure 4. R2DN category and subcategory structure taxonomy

Taxonomy of R2DN Data

As a result of the "Identify disaster data categories" activity represented in Figure 3, we present in Figure 4 a

taxonomy with the structure of identified categories and their subcategories. To construct this taxonomy, we used the DaLA methodology as a reference, which is employed to analyze the impacts generated by disasters on the economy of a society. (World Bank, 2010). The purpose of this representation is to organize the data that will compose the R2DN disaster record, making its understanding and location more accessible. Below, we provide a description of these categories for a better understanding:

- Identification – group all information related to the disaster location, date and time, and event duration.
- Typification –group all information associated with the types and subtypes of the disaster (e.g., hydrological and flooding), disasters associated with the event, and triggering events of the disaster.
- Impacts – human and material impacts caused by the disaster.
- Human – contain the quantity of people directly impacted by the disaster (deaths, missing, injured, etc.).
- Material – group the sectors impacted by the disaster.
- Social – encompass data regarding the quantity of affected housing and educational and health units, whether by destruction or damage, along with associated values for their respective reconstructions.
- Infrastructure – consolidate data related to damages in transportation, water, sanitation, energy, and communication areas.
- Productive – cover damages in production sectors such as agriculture, livestock, fisheries, industry, commerce, and tourism.
- Intersectoral – group damages related to public administration, the financial sector, and the environment.
- Macroeconomic – will contain impacts on inflation, GDP, household income, and public finances.

Table 4. Categories and subcategories and their attributes

Category/subcategory	Attributes
Identification	continent, country, first and second administrative divisions, latitude, longitude, start date and time, duration, name, glide number, source and type of information.
Typification	category, group, subgroup, type, subtype, associated disaster, magnitude value and scale, triggering event, causes and effects.
Humans	dead, injured, sick, missing, displaced, homeless, other affected and people whose livelihoods are affected;
Material impact	description of public and private losses, total public and private losses, total securitized losses and total reconstruction cost;
Impact on housing	destroyed and damaged units, total number of people in destroyed and damaged units and loss value in housing units;
Impact on education	destroyed and damaged units, value to recover the units, value to reestablish teaching and number of days without teaching services;
Impact on health services	destroyed and damaged health units, cost of rebuilding health units, value of reestablishing emergency services, value of loss in vector control and number of days without health services;
Impact on infrastructure	amount of repair and reconstruction works, value of losses for repair and reconstruction, value of losses in public facilities and number of days of interruption in infrastructure services;
Transport	total cost required to restore transport services;
Water and sanitation	losses in the restoration of water supply and restoration of river and sanitary sewage;
Energy	total cost required to restore power generation and fuel distribution;
Communication	total cost to restore telecommunications services;
Agriculture	estimated value of loss of various types of crops;
Livestock	estimated value of losses in dead or sick animals;
Fishery	estimated value of loss in fishing activity;
Industry	estimated value of lost industrial production;
Commerce	estimated value of sales losses from commerce;
Services	estimated value of loss in the services sector;
Other sectors	estimated value of losses in other sectors not previously listed;
Public administration	estimated value to reestablish public safety and urban cleaning services;
Financial sector	estimated value of losses in the financial sector;
Environment	description of environmental damage, population affected by water, air and soil pollution, population affected by water depletion and area affected by fires in parks;
Cultural heritage	estimated value of losses in cultural heritage;

GDP	variation in percentage points of GDP;
Inflation	variation in percentage points in the inflation rate;
Family income	financial impact on family income;
Public finances	financial impact on public resource collection.

Selection of R2DN Attributes And Their Metadata

After analyzing the attributes of each database to address our research question, we adopted the following criteria for selecting the attributes that will be part of the R2DN.

1. **S2ID** – All attributes directly related to disaster events were selected. This selection is based on the R2DN's commitment to providing essential information to the Brazilian federal government to assist states and municipalities affected by a disaster.

2. **EM-DAT** – All attributes related to identification, typification, and impacts associated with the disaster event, as made available on the website, were selected.

3. **DesInventar** – All attributes related to the identification and typification of the disaster were selected. Additionally, among the attributes of human and material impacts, those containing information like S2ID were selected. Also included were those necessary for the indicators of the Sendai Framework.

To meet all the demands of our research question, 101 attributes were chosen, including those from the three selected databases and those necessary for the Sendai Framework indicators. These attributes are distributed across the categories and subcategories presented in Figure 4. In Table 4, we provide a summary of these attributes within their respective subcategories.

The context of this work is the registration of disasters in Brazil. An initiative in another country with the same objectives should encompass items 1 and 2 and adapt item 3 to the context of the new database.

INSTANTIATING R2DN

In this section, we will address the activities necessary to feed the R2DN repository with information about disasters from the selected databases throughout this work. Initially, it is essential to map the R2DN data schema in relation to the schemas of the three databases used in our research: EM-DAT, DesInventar, and S2ID. Integrating disaster records from different databases becomes a valuable tool for comparing different versions of the same disaster.

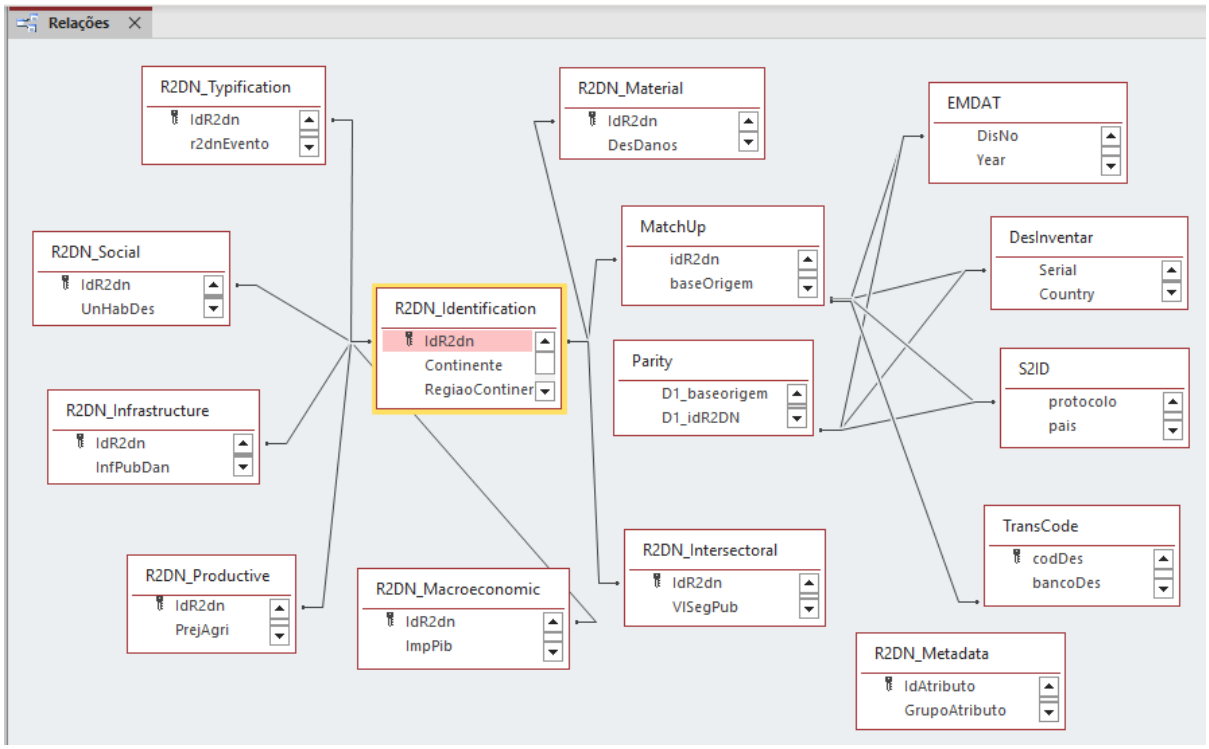


Figure 5. Implementing R2DN in Microsoft Access

Regarding the implementation of the R2DN tables, we decided to organize them based on the data categories established in our taxonomy, as illustrated in Figure 5. In addition to providing a clearer view of the model, this approach was adopted due to the low fill rate in disaster records (CRED, 2021). Therefore, each table will contain records of a disaster only if there is information in the respective data category. The "R2DN_Identification" and "R2DN_Typification" tables, as they contain basic disaster information, will always contain records related to the disaster. Opting for a single-table implementation would have resulted in significant disk space wastage.

The "R2DN_Metadata" table contains the mapping of the attributes of the selected databases with the R2DN attributes. This mapping is necessary to load data from the three databases chosen in R2DN. All other tables with the prefix "R2DN" have information about imported disasters. The additional tables, "MatchUp", "Parity" and "TransCode", play an auxiliary role in the "MatchUp" process, and will be detailed in the "Implementing R2DN" section.

Comparing disasters from different databases can be a challenge, especially when the classifications used are distinct among them. According to (Gall, 2015), it is necessary to establish technical standards for disaster loss databases to ensure comparability and compatibility between databases. The lack of standards hinders the ability to aggregate and consolidate losses over time, space, hazards, and loss metrics, being a critical need in the climate change community. Understanding local, regional, and global conditions is essential for the development of adaptation strategies.

In Table 5 we conducted a comparison of classifications of natural disasters recorded in the databases: EM-DAT (CRED, 2020a), DesInventar (UN Office for Disaster Risk Reduction (UNISDR), n.d.), and S2ID (SEDEC, n.d.). To simplify, we have listed only hydrological and meteorological events, which represent most events in Brazil. Additionally, we limited the presentation to three levels in the classification: Group, Subgroup, and Type.

Table 5. Disaster coding comparison

	EM-DAT		Brazilian disaster code		DesInventar
Group	**Subgroup**	**Type**	**Subgroup**	**Type**	**Type**
Hydrological	Flood	Coastal flood	Flooding		Alluvion
		Riverine flood			Avalanche
		Flash flood			Flood

	Landslide	Avalanche, landslide, mudslide, rockfall and subsidence	Flash flood		Landslide
			Flooding		Rain
Meteorological	Storm	Tropical Cyclone	Large-scale/regional-scale systems	Cyclones	Frost
		Extra-tropical storm		Cold fronts / Convergence zones	Eletricstorm
		Convective storm	Storms	Local / Convective	Snowstorm
	Extreme temperature	Heat wave	Extreme temperature	Heat wave	Heatwave
		Cold wave		Cold wave	Strongwind Tornado
		Severe winter conditions			Hailstorm

Challenges For Data Integration

The challenge in integrating disaster records from different databases lies in identifying similar events across distinct sources. The benefits of identifying similar disasters include:

1. Avoiding distortions in statistical analyses, as some records may refer to the same disaster.

2. Enriching information about a disaster since the three databases contain complementary data.

3. Enabling the comparison of discrepancies in disaster information across various databases.

To identify similarities between disaster records in different databases, three pieces of information are essential: the type of disaster, the location, and the date of the event. The following details how each of these pieces of information will be handled during the R2DN setup process.

As presented in the Table 5, discrepancies in disaster classification among the three databases are noticeable. The difference between S2ID and EM-DAT is relatively smaller, given that the latter based its classification on the former. However, in DesInventar, classification is done by countries, making it more challenging to compare types of disasters. As this work does not aim to exhaust this topic, we choose to restrict the scope to the types of disasters we will address. Thus, we will focus on hydrological, meteorological, and climatological disasters. These categories represent 95.2% of the disasters downloaded from S2ID for the period 1991 to 2022.

Regarding the location of the disaster, greater discrepancies are observed in the information between databases. S2ID considers the location as a discrete domain of states and municipalities, eliminating the possibility of identical locations with distinct spellings. DesInventar also treats the location as a discrete domain, distinguishing administrative divisions by country. For example, in Argentina, it considers province and department, while in Chile, it considers region and province. In EM-DAT, however, the location is a free-text field where various administrative divisions of the country are mixed.

Regarding the date of the disaster, it is necessary to consider a tolerance margin for dates from databases, as there may be a difference of days between dates related to the same disaster in different records. For example, a disaster occurring on 31/12/2013 may be considered the same event in another database but with a date recorded as 01/01/2014. In S2ID, municipalities have a ten-day period to register the disaster in the system from the date of occurrence (UFSC, 2012).

The first activity to instantiate the R2DN involves mapping the proposed data for the R2DN and the data from the selected databases. As the R2DN will be fed with disaster records from the three external databases, it is necessary to identify the source database. From this information, we will use the mapping to find the correspondence between the attributes of the R2DN and the attributes of the source database.

The next step involves generating the files that will contain the disaster records from the three selected databases (CRED, n.d.), (UN Office for Disaster Risk Reduction (UNISDR), n.d.) e (Integração, n.d.), which will be imported into the R2DN. The next section will address the challenges faced when comparing disasters in the R2DN using data from the three databases.

IMPLEMENTING R2DN

The goal of this section is to demonstrate the feasibility of implementing the R2DN through the construction of an artifact, which includes the data structures of the R2DN, the data structures of the three selected databases, and the data structures that will assist in mapping the attributes of each database to the attributes of the R2DN. The main challenge associated with transferring disaster information from other disaster databases to the R2DN is identifying the duplication of disasters. Detecting cases where two or more records from different databases refer to the same disaster is the primary challenge of our experiment, given the disparate ways in which information such as disaster type and disaster location is handled in different databases. This is the challenge we will address during the development of our project.

Since EM-DAT contains disaster records from all countries, DesInventar covers countries in Latin America, excluding Brazil, and S2ID only records events that occurred in Brazil, our goal is to identify duplicate disasters between EM-DAT and DesInventar, as well as between EM-DAT and S2ID. Therefore, it is not expected to find information about an event in all three databases simultaneously.

Regarding the location of the disaster, we note that EM-DAT addresses this information less formally, resulting in a combination of various locations. In the case of Brazil, this may include states and municipalities. On the other hand, S2ID and DesInventar divide the disaster location information into first and second administrative regions, which are pre-registered by countries. Based on this, we conclude that comparing the location in EM-DAT with the other databases should consider both administrative divisions to verify the possibility of similarity of disasters.

Concerning the classification of types of disasters, we observe a significant similarity between the Brazilian disaster code used by S2ID and the disaster classification employed by EM-DAT, as indicated in Table 6. However, the classification used in DesInventar is determined by the country representative when registering the disaster, making it more challenging to find similarities between events. Additionally, we found several instances where the same disaster was registered with different types across databases. For example, as a flood and landslide or a forest fire and drought.

To simplify our observations, we decided to focus exclusively on hydrological, meteorological, and climatological disasters. These categories comprise 95.2% of the disasters recorded in Brazil, as documented in the S2ID Atlas (Integração, n.d.) from 1991 to 2022. Based on our findings, we chose to establish a coding, as illustrated in **Table *6***, for the R2DN. This coding will be used in the comparison between disasters.

Table 6. Disaster Coding Conversion to R2DN

Database	Events	codR2DN
S2ID	Flood, Flash flood, Flooding	1
EMDAT	Flood, Landslide	
DesInventar	Alluvion, Avalanche, Spate	
S2ID	Large scale system, Storm	2
EMDAT	Storm	
DesInventar	Electricstorm, Hailstorm, Snowstorm	
S2ID	Extreme temperature	3
EMDAT	Extreme temperature	
DesInventar	Frost, Heatwave	
S2ID	Seca (Drought)	4
EMDAT	Drought, Wildfire	
DesInventar	Forestfire	

In Figure 6, we present the diagram of the R2DN instantiation process. In the "Load disasters from selected databases" activity, data extracted from the respective databases' websites is imported into auxiliary tables of the R2DN. In the subsequent activity, "Extract data for MatchUp" the "MatchUp" file is generated, containing the data from the three databases necessary for the analysis of similarities between disasters.

The "Analyze similarities between disasters" activity uses the "MatchUp" file to conduct the analysis of disasters. When it identifies one or more similar disasters, it instantiates the "Parity" file with information related to these two disasters. Finally, the "Instantiate R2DN" activity writes the disasters from the three databases into the R2DN.

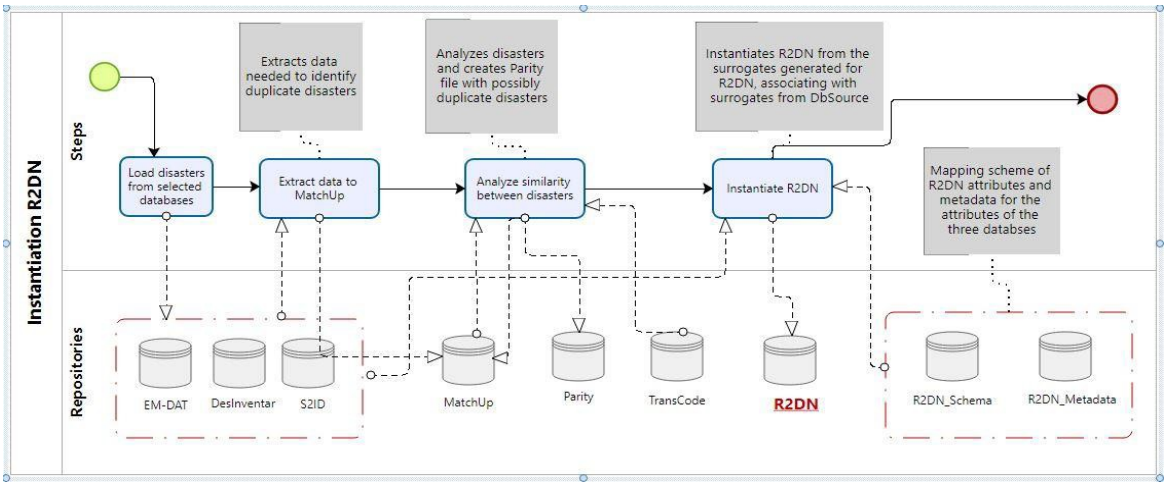


Figure 6. R2DN instantiation process

The "MatchUp" activity, which generates the "Parity" file containing the data of both considered similar disasters, is executed through an SQL command, as illustrated in Figure 7. It is important to note that the "MatchUp" table is invoked twice to enable the comparison of each disaster with all others in search of similarities.

ExecMatch X

```
SELECT d1.baseorigem, d1.idMatch, d1.idbaseorigem, d2.baseorigem, d2.idMatch, d2.idbaseorigem, d1.primDivPais, d2.primDivPais, d1.segDivPais, d2.segDivPais,
d1.data, d2.data, d1.evento, d2.evento, d1.mortos, d2.mortos, d1.feridos, d2.feridos INTO Parity
FROM MatchUp AS D1, MatchUp AS D2
WHERE (D1.pais = D2.pais) AND
    (D1.codEvento = D2.codEvento) AND
    (D1.baseorigem <> D2.baseorigem) AND
    ((D1.contadias - D2.contadias) < 10) AND ((D1.contadias - D2.contadias) > -10) AND
    (((D2.PrimDivPais LIKE '*' & D1.PrimDivPais & '*') OR (D2.SegDivPais LIKE '*' & D1.SegDivPais & '*')) OR
     ((D1.PrimDivPais LIKE '*' & D2.PrimDivPais & '*') OR (D1.SegDivPais LIKE '*' & D2.SegDivPais & '*')));
```

Figure 7. SQL command executed to pinpoint similar disasters

The conditions for identifying similarities, as can be observed in the "WHERE" clause, include: the same country, the same event code, different databases, a date difference of less than ten days, and similarity in the administrative divisions of the disaster locations. Lastly, the " Instantiate R2DN " activity creates disaster records in R2DN from the "MatchUp" file, incorporating the remaining data from the disasters coming from the auxiliary tables of the three selected databases.

The criteria adopted for selecting disasters to be imported from the three selected databases were as follows:

1. Disasters of hydrological, meteorological, and climatological nature.

2. Occurrence in the period from 2010 to 2015.

3. In EM-DAT, only disasters recorded in Brazil, Argentina, and Chile were considered.

4. In DesInventar, only disasters that occurred in Argentina and Chile were included.

The main consideration guiding all the criteria is the quantity of records available in the Brazilian Disaster Atlas Database, accessible through the website http://atlasdigital.mdr.gov.br/paginas/downloads.xhtml (accessed on 10/11/2023). This database covers disasters in Brazil between 1990 and 2022, with a total of 62,273 disaster records. Among these, disasters that fall under the first criterion represent 95.2% of the total.

The disaster records in the DesInventar database for South American countries mostly covered disasters that occurred until 2015. Therefore, we used the criterion of the occurrence period of disasters between 2010 and 2015. The application of this criterion reduced the disasters from the S2ID to 13,143 disaster records.

The choice of Argentina and Chile in South America is due to their presenting similar development indices to Brazil in the region, making the comparison of the impacts of disasters between these countries more equitable. Based on these criteria, the total number of disaster events imported into the auxiliary tables of R2DN was 14,760, as detailed in Table 7.

Table 7. Number of records imported into R2DN

Number of imported disasters in the period 2010-2015	
S2ID	13,143
EM-DAT – Brazil	33
EM-DAT – Argentina	15
EM-DAT – Chile	13
DesInventar – Argentina	1,345
DesInventar – Chile	211
Total	14,760

When considering the criteria for inserting EM-DAT disaster records, which include more than 9 deaths, more than 99 people affected, or the need for international assistance, see Table 8, we assume that all EM-DAT records should be found in the other databases. With this assumption, we face the challenge of locating the 33 EM-DAT disasters that occurred in Brazil in the S2ID, as well as the 28 EM-DAT disasters that occurred in Argentina and Chile in DesInventar.

Table 8. Final result of the MatchUp process

Numbers after running the MatchUp process	
Total records in the Parity file	1,746
Number of similar disasters	873
Similar disasters between EM-DAT and S2ID	686
Similar disasters between EM-DAT and DesInventar	187
Number of disasters not found in S2ID (spelling problem)	3
Number of disasters not found in S2ID	1
Percentage of disasters located in S2ID (except spelling error)	90.63%
Number of disasters not found in DesInventar (spelling problem)	5
Number of disasters not located in DesInventar	7

Percentage of disasters located in DesInventar (except spelling errors)	69.56%

Regarding the 33 Brazilian disasters registered in EM-DAT, one was not identified in our search in S2ID, three were not located due to differences in the spelling of locations, and the remaining 29 show similarities with 686 disasters in S2ID. The way disasters are recorded in S2ID, by municipality, creates situations where, for example, a disaster in the Serrana Region of Rio de Janeiro in 2011, resulting in 900 fatalities, has a single record in EM-DAT, while in S2ID, it is recorded in 19 municipalities.

In another example, the floods that occurred in Rio Grande do Sul in July 2014, due to the overflowing of the Uruguay River, impacted 117 cities, as reported in the media. (G1 - O Globo, 2014). This disaster has only one record in EM-DAT, and our process identified 98 similar disasters in S2ID. This highlights that the information from EM-DAT complements that of S2ID, providing a broader view of disasters, while the latter offers a more detailed perspective on disasters.

Another noteworthy result is that when applying EM-DAT criteria - more than nine fatalities or more than ninety-nine people impacted - for including disasters in its database, we note that, of the 13,143 disaster records in S2ID, 1,924 meet these criteria. According to these parameters, these disasters should be included in the EM-DAT database, which only contains 33 disasters, as shown in **Table *7***. Even considering that one EM-DAT disaster can represent dozens of disasters in S2ID, a considerable disproportion is evident. This reinforces the need for countries to develop regional databases suitable for their realities and demands.

The "MatchUp" process found 90.63% of Brazilian disasters from EM-DAT in S2ID, excluding cases with different spellings. In the case of DesInventar, this percentage was 69.56%. After a preliminary analysis, we were unable to explain the discrepancy between the percentages in the two databases. However, we believe that the greater formality in the use of S2ID may, in part, explain this difference in percentages.

CONCLUSIONS

This study introduced the R2DN data repository model as a response to the research question, which was designed to meet the established requirements. These requirements include the ability to assess disasters, compatibility with major global disaster databases, compliance with multilateral agreements aimed at reducing disaster risks, and meeting the specific needs of a country or organization to support its disaster management policies.

The strategy adopted to achieve the research objectives was as follows:

1. Define a methodological approach for the research.
2. Investigate the characteristics of major global databases, considering citations in academic works.
3. Understand the needs of the Brazilian disaster system, S2ID.
4. Analyze data requirements to meet the demands of the Sendai Framework.
5. Select EM-DAT, DesInventar, and S2ID as reference databases for the study.
6. Identify data needs to meet the requirements of the Sendai Framework.
7. Develop a taxonomy, based on the DaLA methodology, to classify the data selected for the proposed repository model.
8. Implement an artifact to demonstrate the feasibility of the proposed solution.
9. Evaluate the obtained results and generated contributions.
10. Indicate possible directions for future research within the scope of this study.

The **first contribution** of this study lies in the data taxonomy, the data model, and the metadata of R2DN. Our initial encounters with the files of some of the studied databases proved to be quite challenging in understanding the information contained in the data. All databases, without exception, provide their information in a disconnected list of data, significantly complicating its comprehension. Additionally, the available metadata did not offer adequate assistance. We believe that this work will be of great help to those wishing to delve into this subject. The classification of data through the developed taxonomy, coupled with the model that exposes the relationships between the data and the associated metadata, will facilitate the work of anyone dealing with this topic.

The **second contribution** we highlight comes from the development of the artifact that implements an instance of R2DN in a commercial database system. The steps in the development, such as mapping attributes from a database to R2DN, describing data import, unifying disaster type codes, and the procedures necessary to identify similar disasters, facilitate the inclusion of disaster records from other databases into R2DN. We believe we have opened a door for anyone interested in the topic, allowing them to expand R2DN with records from other disaster databases.

Regarding the limitations of this work, we acknowledge that using the three databases restricts the scope of the problem, as the comparison between their data is binary in nature. To elaborate, EM-DAT records need to be compared either with S2ID disasters or with those from DesInventar, given the specificity that S2ID exclusively contains records related to Brazilian disasters, while DesInventar has no records of disasters in Brazil. To overcome this limitation, we intend to include a fourth and perhaps a fifth database in a future version of the work.

Future Work Perspectives

The set of categories does not cover all types of disasters, but it can be easily extended to include disasters such as earthquakes, common in certain parts of the world. As we wanted to examine the occurrence of duplicate events, we left out types of events that are rare in Brazil and therefore are not part of the S2ID database. The taxonomy can also be extended or even replaced by a more complete one.

We believe that handling the location information of the disaster is the aspect that needs the most improvement. The way EM-DAT describes the same disaster occurring in various locations in a single record, with only one attribute for this representation, poses challenges. Furthermore, the spelling of locations in this attribute often contains inaccuracies. Since EM-DAT is the most widely referenced disaster database globally, any initiative to integrate disaster data must include it. In our work, whenever we needed to precisely locate an EM-DAT disaster, we resorted to internet searches using information such as country, date, and type of disaster, and were successful in most cases. Automating this search process represents an opportunity for future work.

Disaster management encompasses three basic phases: prevention, response, and recovery. Our work focused on databases that support the response phase, aiming to restore essential activities, and the recovery phase, to return to pre-event normality in the society affected by the incident. We consider that our proposed disaster data is comprehensive regarding the human and material impacts caused by these events. However, it lacks information on risks in the affected locations, as none of the selected databases in the study address this aspect. Therefore, we believe that integrating a database that includes information on disaster risk into the R2DN represents an opportunity.

Finally, to evaluate our solution, it should be used by researchers and professionals in the field of disasters to demonstrate its efficiency and effectiveness. In our initial attempt to produce such evaluation we didn´t get enough returns to produce a relevant analysis. We intend to pursue this evaluation soon.

ACKNOWLEDGMENTS

The work of Marcos R. S. Borges was partially supported by the Rio de Janeiro Research Foundation (FAPERJ) under grant # E-26/202.876/2018.

REFERENCES

Ceped - Centro Universitário De Estudos E Pesquisas Sobre Desastres. (n.d.). *Base De Dados Do Atlas Digital de Desastres do Brasil (Brazilian Digital Disaster Atlas Database)*. (In Portuguese)

CRED. (n.d.). EM-DAT website. Retrieved from https://public.emdat.be/

CRED. (2020a). *EM-DAT Guidelines : Data Entry , Field Description / Definition*. 1–11.

CRED. (2020b). Human Cost of Disasters. *Human Cost of Disasters*. https://doi.org/10.18356/79b92774-en

CRED. (2021). Missing data on economic losses variables from EM-DAT. *Cred Crunch*, (63), 1–2. Retrieved from https://www.emdat.be/cred-crunch-63-missing-data-economic-losses-variables-em-dat-0

FAPEU. (2020). *Relatório de danos materiais e prejuízos decorrentes de desastres naturais no Brasil : 1995 – 2019 /Banco Mundial.Global Facility for Disaster Reduction and Recovery.* (In Portuguese)

G1 - O Globo. (2014). O Globo. *O Globo*, p. 1. Retrieved from https://g1.globo.com/rs/rio-grande-do-sul/noticia/2014/07/196-mil-pessoas-estao-fora-de-casa-apos-chuva-no-rs-diz-defesa-civil.html

Gall, M. (2015). The suitability of disaster loss databases to measure loss and damage from climate change. *International Journal of Global Warming*, *8*(2), 170–190. https://doi.org/10.1504/IJGW.2015.071966

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly: Management Information Systems*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Integração, M. da. (n.d.). Atlas dos desastres brasileiros (Atlas of Brazilian disasters). Retrieved from http://atlasdigital.mdr.gov.br/paginas/downloads.xhtml (In Portuguese)

IRDR. (2014). *IRDR DATA Reports - Perils Classification and Hazards Glossary*.

Khankeh, S. A. M. F. F. N. N. all 8 authorsHamid R. K. R. (2022). Worldwide disaster loss and damage databases: A systematic review. *Journal of Education and Health Promotion*. https://doi.org/10.4103/jehp.jehp_1525_20

Koç, G., & Thieken, A. H. (2018). The relevance of flood hazards and impacts in Turkey: What can be learned from different disaster loss databases? In *Natural Hazards* (Vol. 91). https://doi.org/10.1007/s11069-017-3134-6

Ministério do Desenvolvimento Regional - S2ID ("Ministry of Regional Development"). (n.d.). Retrieved November 29, 2022, from https://www.gov.br/mdr/pt-br/assuntos/protecao-e-defesa-civil/sistema-integrado-de-informacoes-sobre-desastres#:~:text=O (In Portuguese)

Panwar, V., & Sen, S. (2020). Disaster Damage Records of EM-DAT and DesInventar: A Systematic Comparison. *Economics of Disasters and Climate Change*, *4*(2), 295–317. https://doi.org/10.1007/s41885-019-00052-0

SEDEC. (n.d.). COBRADE. Retrieved from https://www.gov.br/mdr/pt-br/centrais-de-conteudo/publicacoes/protecao-e-defesa-civil-sedec/DOCU_cobrade.pdf

UFSC. (2012). Curso de Capacitação dos Gestores - S2ID ("Manager Training Course"). In *Curso de Capacitação dos Gestores de Defesa Civil para uso do Sistema Integrado de Informações sobre Desastres (S2iD)* . Retrieved from https://s2id.mi.gov.br/

UN. (2015). Sendai Framework 2015-2030. Retrieved from https://www.preventionweb.net/files/43291_sendaiframeworkfordrren.pdf

UN Office for Disaster Risk Reduction (UNISDR). (n.d.). DesInventar Website. Retrieved from https://www.desinventar.net/DesInventar/

UN Office for Disaster Risk Reduction (UNISDR). (2019). *Desinventar Sendai 10.1.2*. (February), 57. Retrieved from https://www.desinventar.net/documentation/DISendai_DataManagement_2019-2.pdf

Wirtz, A., Kron, W., Löw, P., & Steuer, M. (2014). The need for data: Natural disasters and the challenges of database management. *Natural Hazards*, *70*(1), 135–157. https://doi.org/10.1007/s11069-012-0312-4

World Bank. (2010). *Damage, Loss and Needs Assessment: Guidance Notes. Volume 1: Design and Execution of a Damage, Loss and Needs Assessment*. 86. Retrieved from www.worldbank.org

The Role of Digitalization During Supply Chain Disruptions: Lessons from the Experiences of Family-Owned Austrian Firms During the COVID-19 Pandemic

Deepak Khazanchi
University of Nebraska Omaha (USA)
khazanchi@unomaha.edu

Reinhardt Bernsteiner
Management Center Innsbruck (Austria)
reinhard.bernsteiner@mci.edu

ABSTRACT

The COVID-19 pandemic strained global supply chains, and companies faced numerous production, logistics, or material management problems. It is claimed that the digitalization of supply chains can have the potential to make supply chains more resilient to interruptions and disruptions. This article focuses on describing the experiences of three large Austrian family-owned business during the supply chain disruptions caused by the pandemic and the Russia -Ukraine war. This assessment is done from the perspective of the supply chain leaders in the organization. We provide insights on what worked and what did not for these firms and what lessons we can take away from their experiences.

Keywords

Digitalization; Supply Chain Disruptions; Digital Supply Chain; Sustainability; Resilience; Best Practices.

INTRODUCTION

> "*The supply chain mantra is changing from efficiency and low-cost country sourcing, to supply chain resilience and visibility to reduce complexity and uncertainty. The acceleration of digital transformation is key to that new way forward.*" Henry Brunekreef, Director, Operations Advisory and National Leader, Supply Chain Management, KPMG in Australia

The ILO (International Labor Organization) has defined "Global Supply Chains (GSCs)" as the cross-border organization of the activities required to produce goods or services and bringing them to consumers through inputs and various phases of development, production, and delivery. Global supply chains are complex, diverse, fragmented, dynamic, and evolving organizational structures with cross-border production and trade as a critical element (https://www.ilo.org/global/topics/dw4sd/themes/supply-chains/lang--en/index.htm#44).

Established in 2015 by the United Nations, the Sustainable Development Goals (SDGs) are a collection of 17 interlinked goals designed to be a "blueprint to achieve a better and more sustainable future for all". According to (Chauhan et al., 2022; Gunasekaran et al., 2015) the SDG's serve as guidelines for the attainment of sustainable business development and focus on cross-sector as well as cross-country partnerships and cooperation for sustainability. These authors assert that the "efforts to drive businesses towards SDGs are grounded in the equitable sharing of risks and benefits and require actions such as collaborative planning and product development, exchange of information and coordination at various levels among different actors within an GSC" (Chauhan et al., 2022; Mehdikhani & Valmohammadi, 2019). According to the ILO and Pohlmann et al (Pohlmann et al., 2020) to achieve a sustainable future, businesses will need to align their goals with the SDGs at the strategic as well as operational (Chauhan et al., 2022). In fact, Golroudbary et al. (Golroudbary et al., 2019)

have shown in their research that an equal weighting of economic, environmental, and social elements is essential to developing, implementing, and executing sustainable policies and practices. These conclusions are also supported by a study by Haroon et al (Haroon et al., 2021) who explored the perceived significance of sustainability indicators, specifically economic, environmental, and social, among supply chain practitioners in the implementation of sustainable supply chain practices. They discovered that these practitioners consider all three sustainability dimensions as equally important. The SDGs are also well connected with the three pillars of corporate sustainability, ecological, social, and economic, also called the Triple Bottom Line (TBL) (Calabrese et al., 2021). The TBL is frequently employed to translate sustainable development into tangible and actionable decisions (Singh & Rahman, 2021).

Caiado et al. (Caiado et al., 2022) have investigated the integration of SDGs in global supply chain management and report that companies that align their operations and supply chain management activities with the SDGs can positively contribute to:

- ensure reliable and modern energy services, facilitating access to research and technology in clean energy, including renewables, energy efficiency, and advanced and cleaner fossil fuel technology (SDG#7).
- promote sustainable economic growth through the support of productive activities and the creation of decent jobs (SDG#8).
- encourage inclusive and sustainable industrialization by promoting innovation through improved resource-use efficiency and increased adoption of clean and environmentally sound technologies and industrial processes (SDG#9).
- diminish inequality by empowering and fostering the socioeconomic inclusion of everyone while ensuring equal opportunities (SDG#10).
- create more inclusive, safe, resilient, and sustainable cities by establishing accessible, sustainable, and intelligent transport systems (SDG#11).
- enhance technological capacity to transition towards more sustainable consumption and production patterns by minimizing waste generation through prevention, reduction, recycling, and reuse (SDG#12).
- Promote partnerships involving multiple stakeholders to mobilize and share knowledge, expertise, technology, and financial resources (SDG#17).

Supply Chain Disruptions

With the advent of modern digitalization in all its forms, speed, flexibility, scalability, just -in-time inventory, cost-effectiveness, eco-friendliness, or global connectivity are only a few of the many features linked to a global supply chain today. Family-owned businesses that are innovative otherwise often struggle with modernization though digitalization (Baskins, 2023). The COVID-19 pandemic and the Russia -Ukraine war caused supply cha in interruptions and disruptions that demonstrated the vulnerability of an interconnected global economy. Formally, according to Macdonald et al. (Macdonald & Corsi, 2013), supply chain disruptions are defined as "unplanned and unanticipated events that disrupt the normal flow of goods and materials within a supply chain" (Craighead et al., 2007), 132). With COVID-19 (Paul et al., 2021), the Russia-Ukraine war (Dyson et al., 2023), and the recent wildfires across the world (Galizia et al., 2021), it is certain that unanticipated global events will continue to have a significant disruptive impact on global supply chains and, in fact, occur regardless of risk planning (Ivanov & Dolgui, 2021). As a result, supply chain executives continue to face the critical challenge of recovering from disruptions while minimizing their impact.

The COVID-19 pandemic strained global supply chains, and companies faced numerous production, logistics, or material management problems. It is claimed that the digitalization of supply chains can have the potential to make supply chains more resilient to interruptions and disruptions. Some authors have conduced systematic analysis of the academic literature to theoretically discover themes, strategies, and challenges associated with global supply chain disruptions (Hald & Coslugeanu, 2022; Ivanov & Dolgui, 2021; Katsaliaki et al., 2022; Moosavi et al., 2022); there have been no studies that describe lessons and insights from the perspective of large family-owned business and their experiences during the COVID-19 pandemic and other disruptions. There are several researchers who argue that family-owned businesses need to transform particularly during turbulent times (e.g., Davis, 2022), but very little is known about what has worked for them during the disruptions in the past couple of years.

To address the above gap in our understanding, *this paper addresses the following question: What was the impact of digitalization on family-owned businesses during crisis such as Covid-19 and the Ukraine war? This question is explored from the perspective of supply chain leaders in the organization via semi-structured interviews. A descriptive analysis of the interviews is used to elicit the* ***lessons learned*** *relating to the impact of digitalization during supply chain disruptions*.

The remaining paper is organized as follows. Section 2 provides an explanation of our research method followed by a detailed analysis of results in section 3 and a discussion of "lessons learnt" in section 4.

RESEARCH APPROACH

To develop insights into how family-owned firms handled the supply chain disruptions due to the COVID-19 pandemic and the Russia -Ukraine war, we conducted semi-structured interviews over Zoom of senior supply cha in executives from three family-owned Austrian manufacturing firms in Spring 2022. The executives interviewed were the key leaders (SVP or VP level) of the supply chain for each firm and their digitalization efforts. The choice of firms was primarily due to the authors' ability to gain convenient access to the firm and key executives responsible for the digitalization efforts in each firm's supply chain. Although our sample size for this research is a limitation in terms of the generalizability of its conclusions, it is generally challenging, costly and time consuming to obtain access to senior leadership and firms for interviews of the kind reported in this research paper. Furthermore, we concluded that adequate data saturation was achieved with the three firms studied for the research goal of our study. Additional data collection would not produce any value-added insights with regard to the research goal of this research (Adams, 2015b; Saunders et al., 2018). The external validity and generalizability of our conclusions from this sample will of course be limited and applicable to firms that are like those studied. In this context, it is worth reiterating the argument made by Crouch and McKenzie (Crouch & McKenzie, 2006) about the "the logic of small samples…": "in the social world, to discover what features there are in them and to account, however partially, for those features being as they are… a research project scrutinizes the dynamic qualities of a situation (rather than elucidating the proportionate relationships among its constituents), the issue of sample size - as well as representativeness - has little bearing on the project's basic logic."

Expert interviews aim to give further insights about a topic that literature cannot answer entirely (Mayring, 2015b). Although methodological purists raise objections to expert interviews to the expert interview1 on grounds of it being a "dirty" *method,* the value of an expert interview is the individual's viewpoint and experience which can influence the researcher's investigation (Mayring, 2015a; Trinczek, 2009). The open ended semi-structured interviews provided an ideal setting for the free conversational speech of the interviewee combined with guidance from the interviewer (Adams, 2015a, 2015b). To minimize fatigue for both interviewer and respondent, according to (Adams, 2015a, 2015b), about one hour is considered a reasonable maximum length for semi-structured interviews. Our interviews generally lasted 1-1.5 hours with a blend of closed- and open-ended questions, and some follow-up "why" and "how" questions as needed (Adams, 2015a, 2015b).

After each interview, the content was transcribed and anonymized by the authors and coded according to Mayring's scheme for qualitative content analysis (Mayring, 2015a). The authors reviewed each transcript to develop categories or themes from them that address the primary research question of what worked and did not work during the disruptions due to the COVID-19 pandemic and the Russia -Ukraine war and the lessons learnt thereof (Mayring, 2015a).

Table 1: Open ended Semi-Structured Interview Guiding Questions

Q1. Please describe your role and your firm.
Q2. How much digitalization and/or automation exists with respect to the firm's supply chain?
Q3. In the context of the Covid-19 pandemic and the Russia-Ukraine war, how is digitalization and/or digital transformation allowing companies to overcome supply chain disruptions and/or become more resilient?
Q4. What have been the challenges relating to supply chain disruptions in your firm? How is digital transformation helping ("helped")? What are the risks of digitalization?
Q5. How do you (and your firm) think about the resiliency of your global supply chain? What does the future look like for you?

RESULTS

In this section we provide a concise description of the data organized around thematic categories identified using interview evidence and prior research. The three firms studied are characterized in Table 2.

Table 2. Attributes of Firms in this study

Family businesses	Size	Primary Products
Company 1	Employees: 770 Turnover: $730 million. Wholly owned subsidiary of a global company	Manufactures ground-based gardening tools, such as lawn mowers, robotic mowers, ride-on mowers, and garden shredders.
Company 2	Employees: 11,000 Turnover: $2 billion	Truck equipment manufacturer. Makes cranes, hook lifts, cable hoists, forklifts, liftgates, service bodies and platforms. Has 200 suppliers worldwide.
Company 3	Turnover: $1.67 billion Employees: 2,400 worldwide 14 plants; 8 branches in Europe & USA Products available in 100 countries	Business to business provider of beverages including fruit juice production, semi-manufactured products such as concentrates, and branded packaging for business customers.

Qualitative Findings from Interview Data

Nature of Supply Chain Disruptions

The COVID-19 pandemic, Russia-Ukraine war, the cascading impact of China's reaction to COVID-19, transportation bottlenecks, and general geopolitical factors severely impacted global supply chains and to some extent continue to disrupt them. The primary lessons learnt from these challenges to the supply chain are summarized below (Hald & Coslugeanu, 2022; Katsaliaki et al., 2022; Moosavi et al., 2022). These findings combine the information gleaned from prior research and echoed by executives from the three firms studied in this research.

Logistics Disruptions

The global logistics disruptions stemming from the COVID-19 pandemic impacted businesses and consumers as the flow of consumer goods into key markets such as North America and Europe, Southeast Asia and India was restricted by the continued shutdowns of ma jor global ports and airports, largely in China, South Korea, and the USA. According to a supply chain disruptions survey reported by the magazine Logistics Management, as of December 2022, nearly 72% of respondents were still handling global supply chain disruptions and 54% expected this pressure on global supply chains to continue (Bridget McCrea, 2022). The companies we studied faced logistics disruptions relating to semi-conductors or chip bottlenecks from Taiwan. According to the Economist, Taiwan produces over 60% of the world's semiconductors and over 90% of the most advanced ones. This became a limiting factor that intensified competition for chip components in short supply and had a ripple effect on manufacturing during COVID-19 and because of the ongoing the Russia -Ukraine war. All three subject firms faced disruptions due to the Russia -Ukraine war, particularly with respect to semi-conductors or chip bottlenecks impacting firms like Company 1. There was intense competition for chips between the large consumers such as BMW and others in Europe for chip supply). Additionally, the Russia -Ukraine war caused pressure on supply of raw materials such as Aluminum and Wooden Pallets which impacted "Company 1" in particular. Finally, all three subject firms faced transportation and logistics bottlenecks due to lack of shipping containers, time to transport form China/Taiwan to Europe of certain raw materials, and other similar constraints.

Production delays due to lack of components or materials

Production delays during COVID-19 become headline news. Manufactures were competing for limited supply of key commodities and logistical capacity, leading to consumers experiencing empty shelves and long purchase

lead times. This capacity constraint has not completely been ameliorated in many critical sectors. In the technology industry for example, even in 2023, the COVID-19 pandemic, increased demand for consumer electronics, and supply chain disruptions have all contributed to the shortage, which is affecting production, innovation, and prices (Versa Electronics, 2023). The ongoing Russia-Ukraine war has exacerbated the post-COVID recovery by disrupting global supply chains for some basic foods, fertilizers, chemicals, metals, and other minerals as well as oil and gas from Russia (Dyson et al., 2023). As the executive from Company 3 explained: "Out of stock situation are rarely seen in the food industry. However, if the supply chain is disrupted – we know that there are missing materials. Prices go up (supplies, financing, storage, and logistics/transportation) are significant cost factors. If we are out of stock due to bottlenecks – finished goods can be covered for a day or two, but it will ultimately impact production planning. This means we need to be agile while being economically. However, this continues to be a challenge to manage demand planning." For example, even materials that cost a few cents can have a major impact on delivery of materials. Company 3 had to stop production in their brewery p lant due to lack of CO_2 for one week; the cost is a few cents, but it is a critical component and was hard to predict that it would be in short supply. Similarly, Company 3 faced a water shortage due to high temperatures in Hungary being 40C or more; no one imagined such a shortage before. When the Russia -Ukraine war started, Company 3 had to stop receiving nails since steel was restricted from sanctioned countries such as Belarus. A very small item in terms of materia ls needed for the production line, but the intended impact was that palettes to ship goods could not be assembled and their cost tripled.

Over Reliance on limited and remote third parties causing supply chain disruptions

COVID-19 and the Russia-Ukraine war have demonstrated the inherent risk associated with focusing on one major trading partner(s) and/or cheaper remote vendors [(Allam et al., 2022). Many businesses had strong relationships with one major supplier, one large customer and/or one major supply chain partner. Researchers have found that buyers face a tradeoff between the benefits of close relationships with their suppliers, the countries they are located in, and the risk of excessive concentration (Polyviou et al., 2023). For example, according to one study (Oxford Analytica, 2023), Austria had to sharply cut Russia's share of gas imports in the wake of the February 2022 invasion of Ukraine but has already returned to pre-war levels due to challenges associated with contracts, energy diversification, domestic politics and foreign policy. The dependence on China as a manufacturing hub for the world caused server disruptions in the global supply chain during the pandemic and continues to do so even now. In fact it has been reported that nearly 51,000 global companies had one or more tier-1 suppliers, and five million have one or more tier-2 suppliers in the Wuhan region ((Ivanov, 2020); (Jim Kilpatrick, 2020)).

To compensate for the supply chain disruptions brought on by the pandemic and the Russia -Ukraine war, many companies have been moving their manufacturing and/or sourcing closer to home. According to the Logistics Management survey cited previously (Bridget McCrea, 2022), more than half (52.3%) of respondents have attempted to nearshoring and/or reshoring approach over the last two years while 36.4% have not.

In the case of Company 1 and 2 in our study, the reliance on Taiwan for wafers during COVID-19 was a massive limiting factor – demand for chips was high and transportation/logistics bottlenecks were endemic causing delays in production. Furthermore, both were trying to bring some manufacturing of components back to the Eastern Europe where labor was still a bit cheaper. Company 1 was also exploring alternative suppliers for their critical components by trying to just understand which components are critical so that they can move to demand planning better with global standards for suppliers in the value chain. This was driven by worries about not having adequate data transparency for the whole supply chain.

Company 3 being in the food sector had the challenge of understanding the quantity and type of goods that are needed and whether raw materials would be available, particularly since lead time was critical for perishable goods in the food industry. Additionally, dependence on 1000+ farm-based suppliers in Europe was particularly challenging for this organization since many of the farmers have no computers – contracts are manual and individually customized and vary from country to country even within the EU. However, interestingly, billing is electronic while farmers get paper payments.

Technology and Digitalization

The need to improve the resiliency of supply chains has become even more critical post -pandemic. Digita l transformation in supply chain management could enable organizational flexibility, business process automation, cost-effectiveness, visibility, and accelerate innovation in supply chain management itself ((Baryannis et al., 2019), (Hamidu et al., 2023), (Nasiri et al., 2020)). Using digital transformation to this end is a laudable goal but requires financial investment and strategic planning. Absence of planning and collaboration, wrong demand forecast, lack of inventory sharing, incorrect assumptions, lack of knowledge, agility, and flexibility, high volatility, over-reliance on the suppliers, and lack of integration between digital and non -digital management are

some reasons why digital transformation of supply chain may fail ((Baryannis et al., 2019), (Nasiri et al., 2020)). Additionally, there is some argument that digitalization is not a panacea for achieving resilience of global supply chains; the degree and sophistication of the digitalization needed may vary depending upon the nature of the business, size, and concentration of suppliers ((Zavala-Alcívar et al., 2020), (Yin & Ran, 2022), (Shashi et al., 2020), (Casalino et al., 2019)). In fact, Shashi et al. [22] found that "…technology appears as a necessary but not-sufficient enabling factor for ASC ("agile supply chain") deployment."

All the three firms in our study reported varying degrees of success with digitalization before and after COVID - 19 and throughout the disruptions caused by COVID-19 and the ongoing Russia-Ukraine war. At the time of our interview, Company 1 was in the early stages of digital transformation since they had very little plant level automation and used manual processes to handle raw material supplies. They were in the process of implementation of SAP HANA at the plant level with varying adoption across its plant s globally. Even though production planning was automated. Bottleneck handling was mostly annual (prior to COVID-19 through post-COVID timeframe). This was because of lack of data visibility across the supply chain, and it was difficult to get a "clear picture across multiple suppliers." In the words of our interview subject at Company 1, digitalization did not provide any immediate advantage to overcome disruptions, particularly of chips. In fact, Company 1 worked manually first to address manufacturing ba sed on supply of materials – that is adjust what would be manufactured based on what is available from suppliers. They were making decisions about which products to manufacture, for example based on which ones would provide the highest revenue. Limited qua ntity of certain components can impact how to distribute them to the right products in plants. Needed digitalization at the plant level to have all the information relevant for a decision in place, in time and fast.

Company 2 was also in the early stages of digital transformation and was strategically thinking about ways to approach it for its whole supply chain. Their supply chain had different digitalization maturity levels with some plants being the most advanced in terms of VMI, supply relationship management software that connected 200+ suppliers across the globe that also included integration with loading planning and transportation tools. According to our interview subject, Company 1 managed some disruptions in supplies of materials better by using supply relationship management software to get a holistic view of the supply chain. They also used VMI to plan better for inventory. Company 2 also used SAP HANA but for limited functions and was in the process of implementing it across the supply chain.

Company 3 was a contrast to the other two firms since they are in the food services industry and had rarely faced out-of-stock situations until the pandemic. They had implemented EDI (Electronic Data Interchange) with most of their major suppliers. However, they also have 1000+ farmers as fresh commodity suppliers from around Europe who did not have any automation. They had some visibility because of EDI and used demand planning and maintained safety stick when possible. Their mantra was to "be agile while being economical" in what is provided as finished goods. However, demand planning continued to be a challenge. Our interview commented that we because of the way economies are highly interconnected globally, they must use "agile JIT" and not focus on 100% availability of items, which is not always possible as illustrated by the supply chain disruptions during the pandemic and the ongoing Ukraine conflict. They also considered the notion of having more buffers in all parts of the supply chain with reasonable demand planning.

Workforce Issues

Uncertainties in workforce due to shortages of skilled labor and the impact of technology causes challenges for digital transformation of supply chains. To cope with a progressively digitalized and automated supply chain, workers must acquire fresh skills, and advancements in technology will significantly transform their current roles. A considerable portion of employees' routine tasks will be automated, necessitating greater emphasis on innovation, creativity, collaboration, and leadership in their jobs ((Berg et al., 2020), (Foroughi, 2021), (Wahab et al., 2021)). None of the subject companies had workforce concerns at the time of interviews. However, the partners in the supply chain of these firms were globally dispersed and workforce concerns were different in each part of the world. As suggested earlier about the implication of the SDGs on GSCs, there was limited consideration given for the workforce located at supply chain partners in terms of their economic uncertainty created by the pandemic and the impact on the sustainability of their viability. However, Company 3 was particularly concerned about the economic and social impact of the pandemic on the 100s of individual farmers who were their suppliers.

Lack of Transparency/Visibility

Supply chain end-to-end visibility can potentially mitigate the risks of supply chain disruptions (Brandon-Jones et al., 2015; Paul et al., 2021; Roy et al., 2019). Lack of visibility of the supply chain occurs due to inaccurate data, inability to track supplier performance, late updates, and incomplete information. Lack of visibility and transparency can also make decision making challenging in terms of which produ cts and how much to schedule

for manufacturing. Company 1 was in the early stages of digital transformation across the supply chain and lacked an integrated and holistic view of the whole supply chain from sales forecasting to component demand to production planning. In contrast Compa ny 2 used its supply relationship management platform with its 200+ suppliers to gain visibility on information flows and used that to optimize the supplier network particularly during the COVID-19 pandemic Finally, Company 3 was severely challenged with demand planning due to the lack of visibility of the smaller players in their supply chain and due to the unpredictable delays in transportation via containers.

DISCUSSION and LESSONS LEARNED

Based on our analysis of the supply chain disruptions faced by the subject firms and their approaches to addressing them, the following lessons/strategies, and implications for sustainability can be gleaned.

- **Digitalization must be strategically fit(ted)** and it is not a panacea for poor processes or strategic vision. This includes assessing the state of the global supply chain partners in terms of maturity level of digitalization. Strategic digital transformation of GSCs will need huge technology investments and commitment from top management. The implication here is that in terms of sustainability as described in SDG#12 and SDG#8, digitalization across the GSC will involve support for partners to upgrade their infrastructure, train employees, and integrate systems across the supply chain network. For example, Company 1 described the need to responsibly make sure perishable goods are protected and to allow for longer lead times by keeping greater buffer such as an extra truck of supplies. This would need more real-time supply chain optimization.
- Digitalization can result in benefits if the **supply chain automation** includes **much more than just eliminating manual tasks through automation**. It involves process and payment automation, improving resource efficiency, delivering valuable insights, and integrating transportation logistics across the GSC. To reap outstanding rewards, companies need to apply supply chain technology directly within the operations including plant automation and across its whole global supply chain. This insight aligns well with SDG#9 which encourages inclusive and sustainable industrialization by promoting innovation through improved resource-use efficiency and increased adoption of clean and environmentally sound technologies and industrial processes. Company 2 is a great example of how this was taking shape: they had implemented tactical network optimization process and system along with network optimization across their network. They used simulations of their network to conduct scenario analysis. The interview subject commented that "Simulating the network was game changing; the technology was a big advantage." This firm also used their supplier relationship managem ent digital platform to be more resilient by providing clear information flows and had implemented a relationship management software for all its 200 suppliers across globe and were also integrating transportation in the SRM process.
- Digitalization must address the issue of **supply chain transparency and visibility** head on. This not only means clear information flows but also clarification of processes, roles, and responsibilities. This will help with resiliency and adaptability and be in line with SDG#17. The implication of SDG#17 is that true collaboration including sharing of knowledge and information with global supply chain partners is critical for sustainability. We need to work in collaboration with suppliers. As the executive from "Company 2" in our study commented: "… (we) need to change – collaboration with the whole supply chain is essential…. including the internal supply chain – supply chain collaboration in engineering, suppliers, is now needed in the future to be successful".
- Consider **reconfiguring supply chains** to avoid disruption so that there is less dependence on just a few sources of raw materials and/or locations. Need to explore the idea of bringing some critical component manufacturing closer to plant operations. In doing so, the economic and social impac t on supply chain partners across the globe in terms of sustainable practices will need to be addressed per SDG#8 and SDG#17.
- **Adaptability (agility/responsiveness)** of the supply chain through backward integration is essential with digitalization. The idea is that truly resilient companies need to continue processes and make decisions by quickly reacting to dealers, partners, and customers. This means using tools and processes for better demand planning and agility to respond or react to changing demand in collaboration with supply chain partners (SDG#17; SDG#9).
- **Supply chain collaborations** are even more critical for sustainability and resiliency during disruptions (SDG#17; SDG#9). This is particularly important that this happens between large and small suppliers and buyers – VMI by itself cannot address this challenge.
- **Don't forget the human component**; Transparency also means educating the workforce for the next generation of digitalization and process transformations. It also implies that working conditions and

decent work practices be adopted and supported across the GSC network (SDG#8).

Limitations and Implications for Future Research

This research has inherent limitations due to the nature of the qualitative research method and narrow sample of firms participating in the interviews. However, in addition to the lessons learn ed described previously, there are interesting issues that could be further explored using mixed research or survey research designs. For example, research is needed to address the relationship between the extent of digital transformation in a global supply cha in and its impact on reducing supply chain disruptions – this could mean also looking at opportunity costs of not fully digitalizing a supply chain. Additionally, there is still research that needs to be done on how supply chains can be optimally reconfigured to become more resilient through digitalization during disruptions. One of the issues raised by our interviewees was adaptability of the supply chain to disruptions – how can we intentionally design adaptability and responsiveness? Another facet of GSCs that came out in our research that needs further study has to do with engendering effective collaboration across the network due to the large investments needed for a coherent and effective digital transformation strategy. How can firms accomplish this with a shared vision, smart financial investment, governance, and appropriate policies? Finally, a research issue that is often ignored in global supply chains is the study of challenges and best practices of applying decent work principles to labor across the GSC network.

ACKNOWLEDGMENTS

Informed Consent: Interviews were conducted with organizational executives who explicitly agreed to participate in them.

Data Availability Statement: The interview data summarized and presented in this study are available upon request from the corresponding author. The data are not publicly available due to privacy needs of the participating organizations.

Funding: The corresponding author's research time in Austria was made possible by international travel support from a 2022 Fulbright Scholar (Core) Award to Austria.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

Adams, W. C. (2015a). Conducting Semi-Structured Interviews. In *Handbook of Practical Program Evaluation* (pp. 492–505). https://doi.org/10.1002/9781119171386.ch19

Adams, W. C. (2015b). Conducting Semi-Structured Interviews. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of Practical Program Evaluation.* Jossey-Bass. https://doi.org/10.1002/9781119171386.ch19

Allam, Z., Bibri, S. E., & Sharpe, S. A. (2022). The rising impacts of the COVID-19 pandemic and the Russia-Ukraine war: energy transition, climate justice, global inequality, and supply chain disruption. *Resources*, *11*(11), 99.

Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2019). Supply chain risk management and artificial intelligence: state of the art and future research directions. *International Journal of Production Research*, *57*(7), 2179–2202. https://doi.org/10.1080/00207543.2018.1530476

Baskins, K. (2023, February 8). Digital transformation for family-owned business. Available at: https://mitsloan.mit.edu/ideas-made-to-matter/digital-transformation-family-owned-companies.

Berg, L., Farbenblum, B., & Kintominas, A. (2020). Addressing Exploitation in Supply Chains: Is technology a game changer for worker voice? *Anti-Trafficking Review*(14), 47–66. https://doi.org/10.14197/atr.201220144

Brandon-Jones, E., Squire, B., & van Rossenberg, Y. G. T. (2015). The impact of supply base complexity on disruptions and performance: the moderating effects of slack and visibility. *International Journal of Production Research*, *53*(22), 6903–6918.

Bridget McCrea. (2022). *2022 Supply Chain Disruptions Study*. Logistics Management. https://www.logisticsmgmt.com/article/2022_supply_chain_disruptions_study

Caiado, R. G. G., Scavarda, L. F., Azevedo, B. D., Mattos Nascimento, D. L. de, & Quelhas, O. L. G. (2022). Challenges and Benefits of Sustainable Industry 4.0 for Operations and Supply Chain Management—A

Framework Headed toward the 2030 Agenda. *Sustainability*, *14*(2), 830. https://doi.org/10.3390/su14020830

Calabrese, A., Costa, R., Gastaldi, M., Levialdi Ghiron, N., & Villazon Montalvan, R. A. (2021). Implications for Sustainable Development Goals: A framework to assess company disclosure in sustainability reporting. *Journal of Cleaner Production*, *319*, 128624. https://doi.org/10.1016/j.jclepro.2021.128624

Casalino, N., Żuchowski, I., Labrinos, N., Munoz Nieto, Á. L., & Martín, J. A. (2019). Digital Strategies and Organizational Performances of SMEs in the Age of Coronavirus: Balancing Digital Transformation with An Effective Business Resilience. *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.3563426

Chauhan, C., Kaur, P., Arrawatia, R., Ractham, P., & Dhir, A. (2022). Supply chain collaboration and sustainable development goals (SDGs). Teamwork makes achieving SDGs dream work. *Journal of Business Research*, *147*, 290–307. https://doi.org/10.1016/j.jbusres.2022.03.044 .

Craighead, C. W., Blackhurst, J., Rungtusanatham, M. J., & Handfield, R. B. (2007). The severity of supply chain disruptions: design characteristics and mitigation capabilities. *Decision Sciences*, *38*(1), 131–156.

Crouch, M., & McKenzie, H. (2006). The logic of small samples in interview-based qualitative research. *Social Science Information*, *45*(4), 483–499. https://doi.org/10.1177/0539018406069584

Davis, J. (2022). Transform a family business during turbulent times. *Harvard Business Review*. Available at: https://hbr.org/2022/09/transforming-a-family-business-in-turbulent-times.

Dyson, E., Helbig, R., Avermaete, T., Halliwell, K., Calder, P. C., Brown, L. R., Ingram, J., Popping, B., Verhagen, H., Boobis, A. R., Guelinckx, I., Dye, L., & Boyle, N. (2023). Impacts of the Ukraine–Russia Conflict on the Global Food Supply Chain and Building Future Resilience. *EuroChoices*, *22*(1), 14–19. https://doi.org/10.1111/1746-692X.12380.

Foroughi, A. (2021). Supply chain workforce training: addressing the digital skills gap. *Higher Education, Skills and Work-Based Learning*, *11*(3), 683–696. https://doi.org/10.1108/HESWBL-07-2020-0159

Galizia, L. F., Alcasena, F., Prata, G., & Rodrigues, M. (2021). Assessing expected economic losses from wildfires in eucalypt plantations of western Brazil. *Forest Policy and Economics*, *125*, 102405.

Golroudbary, S. R., Zahraee, S. M., Awan, U., & Kraslawski, A. (2019). Sustainable Operations Management in Logistics Using Simulations and Modelling: A Framework for Decision Making in Delivery Management. *Procedia Manufacturing*, *30*, 627–634. https://doi.org/10.1016/j.promfg.2019.02.088

Gunasekaran, A., Subramanian, N., & Rahman, S. (2015). Green supply chain collaboration and incentives: Current trends and future directions. *Transportation Research Part E: Logistics and Transportation Review*, *74*, 1–10. https://doi.org/10.1016/j.tre.2015.01.002

Hald, K. S., & Coslugeanu, P. (2022). The preliminary supply chain lessons of the COVID-19 disruption—What is the role of digital technologies? *Operations Management Research*, *15*(1), 282–297. https://doi.org/10.1007/s12063-021-00207-x

Hamidu, Z., Mensah, B. D., Issau, K., & Asafo-Adjei, E. (2023). Does technological innovation matter in the nexus between supply chain resilience and performance of manufacturing firms in a developing economy? *Journal of Manufacturing Technology Management* .

Haroon, S., Wasif, M., Khalid, R., & Khalidi, S. (2021). Supply Chain Practitioners' Perception on Sustainability: An Empirical Study. *Sustainability*, *13*(17), 9872. https://doi.org/10.3390/su13179872

Ivanov, D. (2020). Predicting the impacts of epidemic outbreaks on global supply chains: A simulation -based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. *Transportation Research Part E: Logistics and Transportation Review*, *136*, 101922.

Ivanov, D., & Dolgui, A. (2021). A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. *Production Planning & Control*, *32*(9), 775–788. https://doi.org/10.1080/09537287.2020.1768450

Jim Kilpatrick. (2020). *Managing supply chain risk and disruption.* Deloitte Canada. https://www.deloitte.com/global/en/services/risk-advisory/analysis/covid-19-managing-supply-chain-risk-and-disruption.html

Katsaliaki, K., Galetsi, P., & Kumar, S. (2022). Supply chain disruptions and resilience: a major review and future research agenda. *Annals of Operations Research*, *319*(1), 965–1002. https://doi.org/10.1007/s10479-020-03912-1

Macdonald, J. R., & Corsi, T. M. (2013). Supply Chain Disruption Management: Severe Events, Recovery, and Performance. *Journal of Business Logistics*, *34*(4), 270–288. https://doi.org/10.1111/jbl.12026

Mayring, P. (2015a). Qualitative content analysis: Theoretical background and procedures. *Approaches to Qualitative Research in Mathematics Education: Examples of Methodology and Methods*, 365–380.

Mayring, P. (2015b). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., überarb. Aufl.). *Beltz Pädagogik*. Beltz.

Mehdikhani, R., & Valmohammadi, C. (2019). Strategic collaboration and sustainable supply chain management. *Journal of Enterprise Information Management*, *32*(5), 778–806. https://doi.org/10.1108/JEIM-07-2018-0166

Moosavi, J., Fathollahi-Fard, A. M., & Dulebenets, M. A. (2022). Supply chain disruption during the COVID-19 pandemic: Recognizing potential disruption management strategies. *International Journal of Disaster Risk Reduction*, *75*, 102983. https://doi.org/10.1016/j.ijdrr.2022.102983

Nasiri, M., Ukko, J., Saunila, M., & Rantala, T. (2020). Managing the digital supply chain: The role of smart technologies. *Technovation*, *96-97*, 102121. https://doi.org/10.1016/j.technovation.2020.102121

Oxford Analytica (2023). Austrian exit from Russian gas is far from assured. *Emerald Expert Briefings*(oxan-db).

Paul, S. K., Chowdhury, P., Moktadir, M. A., & Lau, K. H. (2021). Supply chain recovery challenges in the wake of COVID-19 pandemic. *Journal of Business Research*, *136*, 316–329. https://doi.org/10.1016/j.jbusres.2021.07.056

Pohlmann, C. R., Scavarda, A. J., Alves, M. B., & Korzenowski, A. L. (2020). The role of the focal company in sustainable development goals: A Brazilian food poultry supply chain case study. *Journal of Cleaner Production*, *245*, 118798. https://doi.org/10.1016/j.jclepro.2019.118798

Polyviou, M., Wiedmer, R., Chae, S., Rogers, Z. S., & Mena, C. (2023). To concentrate or to diversify the supply base? Implications from the US apparel supply chain during the COVID-19 pandemic. *Journal of Business Logistics*.

Roy, A., Gilbert, S. M., & Lai, G. (2019). The implications of visibility on the use of strategic inventory in a supply chain. *Management Science*, *65*(4), 1752–1767.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135 -017-0574-8

Shashi, Piera Centobelli, Roberto Cerchione, & Myriam Ertz (2020). Agile supply chain management: where did it come from and where will it go in the era of digital transformation? *Industrial Marketing Management*, *90*, 324–345. https://doi.org/10.1016/j.indmarman.2020.07.011

Singh, A. P., & Rahman, Z. (2021). Integrating corporate sustainability and sustainable development goals: towards a multi-stakeholder framework. *Cogent Business & Management*, *8*(1), Article 1985686. https://doi.org/10.1080/23311975.2021.1985686

Trinczek, R. (2009). How to interview managers? Methodical and methodological aspects of expert interviews as a qualitative method in empirical social research. In A. Bogner, B. L. Littig, & W. Menz (Eds.), *ECPR Research Methods. Interviewing experts* (pp. 203–216). Palgrave Macmillan.

Versa Electronics. (2023). *Electronic Component Shortage – An Update for 2023*. https://versae.com/electronic-component-shortage-2023-update/

Wahab, S. N., Rajendran, S. D., & Yeap, S. P. (2021). Upskilling and reskilling requirement in logistics and supply chain industry for the fourth industrial revolution. *Logforum*, *17*(3), 399–410. https://doi.org/10.17270/J.LOG.2021.606

Yin, W., & Ran, W. (2022). Supply chain diversification, digital transformation, and supply chain resilience: Configuration analysis based on FSQCA. *Sustainability*, *14*(13), 7690.

Zavala-Alcívar, A., Verdecho, M.-J., & Alfaro-Saiz, J.-J. (2020). A conceptual framework to manage resilience and increase sustainability in the supply chain. *Sustainability*, *12*(16), 6300.

Comparative Analysis of Strategies for National Food Stockpiling: A Case Study of Germany and Switzerland

Katharina Eberhardt*
Institute for Industrial Production (IIP),
Karlsruhe Institute of Technology (KIT)
katharina.eberhardt@kit.edu

Simon Stieler
Institute for Industrial Production (IIP),
Karlsruhe Institute of Technology (KIT)

Florian Klaus Kaiser
Institute for Industrial Production (IIP),
Karlsruhe Institute of Technology (KIT)

Frank Schultmann
Institute for Industrial Production (IIP),
Karlsruhe Institute of Technology (KIT)

ABSTRACT

Effective disaster preparedness, including the strategic stockpiling of resources, is crucial to respond quickly and efficiently to unforeseen crises. Therefore, we contribute to understanding effective disaster preparedness by examining the national food stockpiling practices in Germany and Switzerland, nations renowned for their proactive disaster management and high levels of preparedness. Using comparative content analysis, we thoroughly investigate the strategies, stakeholder involvement, and economic aspects these nations employ to ensure food supply. Insights are gained through in-depth interviews with experts in disaster management, food production, and food supply. The analysis reveals significant differences in the approaches, highlighting benefits and drawbacks for policymakers and practitioners. Our research underscores the importance of different strategies to suit national contexts and provides insights into best practices for enhancing food security and supply in the face of disasters.

Keywords

Preparedness, Expert interview, Disaster management, Decision-support

ACKNOWLEDGMENTS

This work/the project is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE), grant number 2821HS012.

INTRODUCTION AND MOTIVATION

In today's ever-changing world, crises like the 2021 century flood in Germany, the COVID-19 pandemic, and military conflicts consistently impact global supply chains, particularly in the food sector. Given its intricate nature and essential role, food supply chains face an extensive risk landscape due to several factors. Notably, supply chain disruptions are increasing due to the growing frequency of crises and natural disasters, accompanied by changing climate patterns (WEF, 2023). Furthermore, the expanding global population drives increased demand for food, placing significant pressure on supply chains to ensure efficient and timely delivery (Rejeb et al., 2022). Additionally, risks arise from global food supply networks, technological dependencies, regulatory complexities, and growing interconnectedness (Kumar et al., 2022). The COVID-19 pandemic highlighted these vulnerabilities and significantly impacted and disrupted food supply chains. Bottlenecks in production and transport led to a mismatch between demand and supply. They impacted the availability and access to certain food products as described by Mor et al. (2021) and Weersink et al. (2021). The food supply chain relies on a complex and highly

*corresponding author

interconnected structure across various industries. Disruption or failure in any part of this chain can lead to systemic collapse. Dependencies within the food supply chain extend to critical sectors such as energy, water, transport or communication (Brem, 2015). These interdependencies increase the vulnerability of the food sector, with the extent of their impact varying based on the crisis.

As a result, governments, organizations, and private actors are urged to invest in strengthening food production capacities and establishing contingency plans that decision-makers can swiftly activate in times of crisis. One strategy in this regard is the pre-positioning of emergency reserves. For example, Germany and Switzerland have adopted measures to stockpile strategic food reserves and promote collaboration among stakeholders (BMEL, 2023). It is essential to note that both countries' disaster preparedness extends beyond food supplies, acknowledging that crises impact multiple sectors and require diverse relief materials such as energy and healthcare resources (BMWK, 2024; FONES, 2023). Despite their interconnectedness, sectors are usually managed independently, emphasizing unique stockpiling requirements. Notably, storing food presents distinct challenges, including meeting nutrition needs and changing population demands depending on the type of crisis.

Therefore, our paper examines the crucial significance and approaches of food stockpiling. It investigates emergency preparedness and storage strategies in Germany and Switzerland through four expert interviews, gathering valuable insights from professionals with expertise in the field. Based on the focus group interviews, this work provides insights into the following research questions:

- RQ1: What are the distinctions in stockpiling strategies between Germany and Switzerland, considering factors such as stakeholder involvement and economic aspects?
- RQ2: What insights can be derived from the national strategies to enhance preparedness and optimize the efficient distribution of resources in disaster preparedness?

By answering these research questions, our study provides the following contributions: First, by conducting and analyzing four expert interviews, we aim to identify the strategies, motivations, and decision-making processes involved in implementing food reserves in disaster preparedness. This approach allows us to gain valuable insights from experts with practical experience in different sectors and countries. Second, we thoroughly examine the advantages and challenges of the presented food stockpiling strategies, providing a comprehensive perspective on the efficiency and feasibility. Third, utilizing the insights gained from the expert interviews and the analysis of stockpiling strategies, we aim to provide stakeholders with a comprehensive understanding of different approaches to food emergency preparedness, aiding stakeholders in making well-informed and resilient decisions that can shape policies and actions. Fourth, in the discussion section, we highlight the critical importance of adapting to different crises to ensure the effectiveness of systems in diverse settings. These considerations enable decision-makers to tailor their strategies and responses accordingly. Furthermore, we offer recommendations for future work, focusing on understanding the nuances of different crises and leveraging tools such as information systems to manage the resulting complexities and enhance overall crisis resilience by taking on an integrated view.

This study is structured as follows. Initially, a comprehensive overview of the theoretical background is provided, setting the foundation for the subsequent discussion. Next, the research methodology is outlined, detailing the approach for gathering and analyzing data. Subsequently, the research outcomes are presented along with the critical findings. Finally, the study discusses the results and explores potential opportunities for further research.

THEORETICAL BACKGROUND

Pre-Positioning of Relief Supplies and National Stockpiles

Pre-positioning relief supplies is critical in disaster management to strengthen resilience and response operations. It involves strategically placing relief supplies such as food, water, medical supplies, and other essential commodities before a crisis occurs (Sabbaghtorkan et al., 2020).

Numerous studies, including those by Arnette and Zobel (2019), Condeixa et al. (2017), Ni et al. (2018), Rawls and Turnquist (2010), and Verma and Gaukler (2015) have explored the importance and implementation of strategies for stockpiling critical goods. Ongoing research in food emergency preparedness focuses on assessing the unique advantages of collaborative efforts in food supply during emergencies (Edge & Meyer, 2019). These works include studies aiming to enhance the efficiency of logistical processes and ensure the timely distribution of food resources (Ö zdamar et al., 2004). Furthermore, regulatory frameworks (Manzoor et al., 2014) and technological collaborations (Shrivastav & Bag, 2023) are investigated.

Several countries maintain national stockpiles, incorporating critical goods and food items in line with these considerations. The objective is to safeguard the food supply during crises, encompassing scenarios like natural disasters, wars, or other supply disruptions. The composition and scale of strategic national food stockpiles vary between countries depending on the country's emergency strategy and factors such as self-sufficiency, population size, geographical location, and vulnerability (Caballero et al., 2015). For example, Germany and Switzerland stockpile food for crises, while other countries, such as France or Italy, forego strategic national food reserves (Gerhold et al., 2019).

While the stockpiling of supplies in humanitarian relief operations is a subject of intense discussion, there remains a notable knowledge gap in comprehending how specific strategies can enhance pre-positioning and logistical activities, especially in the context of food-related crises at the national level.

Stockpiling Policies in Germany and Switzerland

German law mandates state emergency food provisions as part of the broader obligation to safeguard every human life (BVerfG, 1977). The Federal Ministry of Food and Agriculture (BMEL) is the primary entity responsible for emergency food provision in Germany. It is mandated to enact sovereign food management during a crisis (BMEL, 2023). Acting on behalf of the BMEL, the Federal Agency for Agriculture and Food (BLE) maintains about 150 warehouses known as the Civil Emergency Reserve and the Federal Grain Reserve, storing essential items such as rice, lentils, condensed milk, wheat, rye, and oats (BLE, 2023b). These stored provisions are intended to ensure a consistent supply of flour and bread and sustain the overall population's needs in urban areas (BMEL, 2023). Over the years, Germany has significantly transformed its approach to food security. The country incorporated military provisions into the constitutional framework in 1956, laying the groundwork for emergency food supply (Menski, 2016). Motivated by Cold War tensions, Germany instituted the Food Security Act (ESG) in 1965 and, in response to the Chornobyl disaster in 1986, introduced the Food Precaution Act (EVG) in 1990 (BLE, 2023a; Menski, 2016). These acts ensure essential provisions for the civilian population and the armed forces during war and other crises. In 2017, these acts were consolidated into the Food Supply and Preparedness Act (ESVG), providing a unified approach to defense and non-defense crises (BLE, 2023a).

In Switzerland, the national economic supply provides the country with essential goods and services in times of shortages (FONES, 2023). The National Economic Supply System (NES) is based on cooperation between businesses and the state. Representatives from the key economic sectors take on management responsibility and contribute their experience (FONES, 2021). Companies that import, manufacture, process, or introduce essential goods for the first time into the market must maintain reserves (Re´servesuisse, 2023). Storage requirements for food and feed, including items like sugar, rice, edible oils and fats, coffee, and grains for human consumption, as well as energy and protein carriers for feed purposes, totaling 711,294 tons (FONES, 2022, 2023). The Federal Office for Economic National Supply (FONES), which is subordinate to the Federal Department of Economic Affairs, Education and Research (EAER), is responsible for supervising the process (FONES, 2021). The supply strategy generally consists of two main phases: the preparatory phase and the intervention phase during severe shortages (FONES, 2018). In the preparatory phase, resilience is strengthened through training and collaboration with stakeholders, businesses, and industry associations to develop measures for improving resilience and interventions. According to the FONES (2018), the intervention phase comprises three stages:

1. Stage A (severe shortage less than one year): Stock releases, import promotions, and export restrictions.
2. Stage B (severe shortage up to one year): Stock releases and supply restrictions to distribute the scarce goods evenly to the population.
3. Stage C (severe shortage lasting over a year): Switch to domestic production to ensure a supply of 2,300 calories per person per day.

As the severity of the shortage situation improves, interventions are gradually withdrawn, aiming to return to normal conditions as soon as possible (FONES, 2022). Switzerland's supply policy historically transformed from state interventionism to market-driven approaches in response to changing global dynamics (FONES, 2011). During the First World War, the Swiss Federal Council slowly recognized the importance of food security and strategically intervened in the market to supply the population in crisis (Re´servesuisse, 2023). From 1929, companies were first obligated to store grain as part of the non-monopoly grain supply system, a hybrid of free-market economy and state-controlled areas (Willisegger, 2015). In 1955, the mandatory storage of these reserves was more comprehensively regulated by the Federal Act on Economic Preparedness for Wartime (FONES, 2011). 1982 the legal basis was established, allowing the state to intervene subsidiarily even in severe shortages outside of war

(Willisegger, 2015). This shift was prompted by the realization that supply threats could extend beyond wartime scenarios, as demonstrated by events such as the Suez Crisis and OPEC's actions in 1973 (Rickenbacher, 2021; SWI, 2022). This gradual transition led to the Federal Act on National Economic Supply (NESA), which expanded the scope of economic national defense (SE, 1982). Under this act, the federal government can implement state supply measures during war, power-political threats, and scarcity. On March 30, 2022, the Federal Council strengthened further supply measures in response to new requirements influenced by the energy crisis (SE, 2023). The detailed historical development of both stockpiling policies is depicted in Figure 1.



Figure 1. Historical development of stockpiling policies in Germany and Switzerland

METHODOLOGY

Case Description

We selected Germany and Switzerland as case-study countries for food emergency preparedness to further analyze the importance of pre-positioning and the connection between the public and private sectors. The approach offers a comprehensive perspective on various aspects of food security, ranging from economic significance and different governance models to internal and external collaboration.

Germany and Switzerland have established measures for national food stockpiling to prepare for potential crises. However, the approaches differ significantly. While Germany pursues a state-focused approach, Switzerland delegates significant portions of the responsibility to the private sector. Given the notable differences in the approaches of these neighboring nations, it is interesting to highlight their distinct development and strategies and present their advantages and disadvantages, along with the varying viewpoints.

Study Design

A qualitative study was conducted based on in-depth semi-structured interviews with experts in Germany and Switzerland. Expert interviews were chosen as they provide valuable perspectives from professionals specializing in food stockpiling, encompassing food security, logistics, and disaster management. The semi-structured interview method allows interviewees to freely articulate their thoughts on the subject matter, facilitating an open dialogue that can yield important insights beyond the constraints of a predefined question format.

Data Collection

A purposive sampling approach was employed to ensure the inclusion of experts actively engaged in the food supply chain. The sample criteria involved reaching out to individuals from governmental organizations and private companies with institutional and practical knowledge of emergency procedures, strategies, and agencies within the food supply chain. Initial contact was made via email or phone, introducing the research objectives and inviting

them to participate. Participants from Germany and Switzerland were explicitly addressed in the study to ensure comparability between the food emergency systems. However, finding participants proved to be highly challenging due to the sensitive nature of the topic and the limited engagement of the private sector in Germany.

In total, four interviews were conducted during July and August 2023. The selected experts are widely acknowledged for their extensive understanding of the topic within their respective countries, and their expertise is derived from their specific roles and responsibilities, as described in Table 1. The Semi-structured interviews were conducted via MS Teams, using separate interview guides for two stakeholder groups: i) Public authorities (policymakers) and ii) Private companies in the food industry (retailers and producers). The interviews, conducted in German, lasted between 60 to 90 minutes. All interviews were recorded with participant consent and transcribed verbatim for analysis.

Code	Role	Type of Organization		Country	
		Public	Private	DE	CH
DE1	Federal Agency for Agriculture and Food (BLE)	X		X	
DE2	Branch manager of several supermarkets		X	X	
CH1	Federal Office for Economic National Supply (FONES)	X			X
CH2	Head of procurement of a milling company		X		X

Table 1. Overview of expert interviews on food emergency stockpiling

Data Analysis

For thematic data analysis, we employed a qualitative content analysis approach based on Mayring and Gahleitner (2010). The statements obtained from the interviews were systematically structured into categories and coded to identify emerging themes and sub-themes. These provided valuable information about various aspects such as strategies, stakeholders, economic processes, and perceived advantages and disadvantages. To ensure the reliability of findings, different researchers coded the contents independently. After that, the results were compared and discussed until a consensus on the themes and sub-themes was achieved. This collaborative process resulted in a comprehensive coding framework that served as a basis for further interpretation and understanding of the data, as shown in Table 2.

Categories	Subcategories
Stockpiling strategy	Structure
	Development and significance
Stakeholder	Responsibilities
	Collaboration
Economic aspects	Financing
	Costs
System comparison	Strategy
	Stakeholder
	Economic aspects

Table 2. Categorization of interview data

RESULTS: COMPARATIVE ANALYSIS OF FOOD RESERVES IN GERMANY AND SWITZERLAND

Stockpiling Strategy

Germany and Switzerland adopt distinct approaches in implementing their food emergency preparedness systems, although some similarities exist. Both nations prioritize accumulating essential foods like grains due to their convenient storage and prolonged shelf life compared to processed products. Nevertheless, a direct comparison highlights that Switzerland also includes edible oils and coffee in its reserves. The duration for which the state's reserves can sustain the German population during a crisis remains to be determined as no legal supply period is specified. In contrast, stock levels in Switzerland are closely integrated into the regular economic production process and are set at three to four months. Calculations are grounded in import statistics and domestic production over a four-year cycle, automatically adapting to shifts such as immigration or population growth.

According to the expert *DE1*, Germany maintains emergency food reserves and other national reserves to reduce dependence on the food retail sector and fulfill the government's mission to supply the population during crises (DE1,

p. 20, lines 39-45). This state-centric strategy aims to strengthen the national emergency food supply, consistently maintain predetermined quantities of specific food items, and release them in crises. The reserves are allocated across several storage facilities, including the Federal Grain Reserve and the Civil Emergency Reserve, strategically positioned at confidential sites to prevent looting (DE1, p. 4, lines 16-18). The warehouses are contracted to private companies through tenders for storage and related activities, while ownership of the goods and control remains with the government (DE1, p. 20, lines 8-9). In a supply crisis, the federal states can apply for access to goods from the warehouses. The processing, transportation, and distribution of these goods are then the responsibility of the requesting federal state (DE1, p. 15, lines 1-5).

DE1 underscores that initially, the government aimed to avoid entrusting the protective role of food supply to private entities (DE1, p. 20, lines 39-44). This foundational position shifted during the 1990s with a broader reconsideration of the necessity of stockpiling and the government's role. Consequently, by 1995, private entities were operating the warehouses on behalf of the state (DE1, p. 15, lines 39-42). According to *DE1*, this shift is connected to the growing perception that stockpiling is no longer essential, a sentiment that gained momentum during that period (DE1, p. 20, lines 3-15). The overall attitude towards stockpiling has undergone multiple transformations since the conclusion of both World Wars. *DE1* emphasizes that *"there used to be a seriousness"*, backed by the organizational structure (DE1, p. 6, lines 9-10). Formerly, three departments oversaw food stockpiling, whereas today, only one remains (DE1, p. 6, lines 10-15). The peak of the contrasting attitude towards stockpiling occurred in 2016 when the Federal Minister of Agriculture contemplated the complete abolition of emergency food supply (DE1, p. 4, lines 38-44). However, according to *DE1*'s observations, the general attitude has shifted in the opposite direction in recent years. The Ukraine-Russia war has strengthened the current stance where the department overseeing stockpiling is directed to take all necessary measures to enhance supply security (DE1, p. 4, lines 42-44). This change is also emphasized by *DE2*, who deems preparatory measures crucial to mitigate the impact of unforeseen crises (DE2, p. 7, lines 6-14). Despite the significance, *DE1* contends that *"stockpiling has been overlooked in the past"* (DE1, p. 21, line 10), and the awareness of the topic is influenced by global market dynamics and worldwide crises (DE1, p. 21, lines 6-11).

In Switzerland, the State mandates the stockpiling of essential goods, imposing mandatory reserves, the so-called National Economic Supply System (NES). According to *CH1*, specific market players are legally obligated to maintain reserves for three to four months of their typical product processing or importing (CH1, p. 3, lines 42-45). Thereby, the EAER determines the duration for which these mandatory reserves must cover the average demand of the Swiss population (CH1, p. 17, lines 14-16). These mandatory reserves, integrated into companies' production sites, aim for efficient turnover during crises. Approximately 60 companies manage mandatory stocks of essential goods, such as sugar, wheat, rice, oils, fats, yeast, and coffee (CH1, p. 3, lines 45-52 & p. 6, lines 24-25).

According to *CH2*, the organization Re`servesuisse calculates mandatory stock levels for each company on behalf of the FONES based on import volumes and domestic production (CH2, p. 7, lines 39-45). A comprehensive review is conducted every four years, adjusting compulsory stock quantities for all products as needed (CH2, p. 8, lines 3-7). Notably, the stocked quantities remain private and not state-owned. In severe shortages, these stocks can be released to prevent or mitigate disruptions in the supply of essential goods (CH1, p. 14, lines 9-10). *CH2* clarifies that Switzerland's stockpiling strategy is not a response to individual planning failures of specific market participants but rather a macroeconomic assessment to identify and address market-wide crises, determining the need for stock releases (CH2, p. 5, lines 40-49).

In the early stages of the Swiss emergency food system, expert *CH1* notes that the State exclusively managed grain storage from the First World War to the 1970s and 1980s. The Swiss government collected harvested grain at designated points and oversaw its administration (CH1, p. 15, lines 21-23). A shift in the 1980s, driven by economic changes and a decline in self-sufficiency, led to increased imports and a transfer of food supply responsibility to the private sector under state supervision (CH1, p. 15, lines 31-45).

CH1 notes a significant change in Switzerland's perspective on emergency food stockpiling in recent years. Five years ago, many politicians and the population considered the provision system outdated and old-fashioned (CH1, p. 25, lines 49-52). However, the mentality changed with the outbreak of the COVID-19 pandemic in 2020:

> *"It was not necessarily the pandemic itself, but the disruptions that occurred in high-risk supply chains that led to a rethink of the need for reliable emergency food stockpiling."*
> (CH1, p. 16, lines 1-3)

This rethink was further reinforced when the Ukraine-Russia war highlighted the fragility of global supply chains. Today, the focus is no longer on abolishing emergency food preparedness but on analyzing where additional measures and resources are needed (CH1, p. 16, lines 1-13).

Regarding the historical evolution of food emergency preparedness, Germany and Switzerland initially shared common approaches but diverged significantly over time. Both nations maintained complete state control over warehouses and crisis supplies during the World Wars. In the 1980s, Switzerland shifted responsibility to the private sector due to a declining farming population and increased import dependence. Germany underwent a similar change in the 1990s, privatizing warehouses as the perceived necessity of state-controlled food storage diminished.

Today, Germany's strategy still centers mainly on the state's provision of food reserves, aligning with constitutional responsibilities for meeting basic needs. In comparison, Switzerland adopts a collaborative approach between the federal government and businesses. According to the law, companies are primarily responsible for supplying the population and must maintain mandatory stockpiles covering three to four months. The organization Re`servesuisse manages quantity calculations and stock allocations, with private companies owning warehouses and stored goods. In Germany, private companies secure storage spaces through tenders, while the state retains ownership of the stored goods. The quantities stored lack a legal foundation but are based on a 2011 decree that has yet to undergo fundamental adjustments in response to societal changes. Hence, in contrast to Switzerland, quantity requirements are determined by a range rather than a timeframe for supply.

In summary, the two countries exhibit different approaches in their current strategies, diverging from their earlier similarities. Germany emphasizes state control and provision, while Switzerland increasingly relies on collaboration with the private sector. Both nations respond to global events, demonstrating heightened awareness of the importance of food emergency preparedness.

Stakeholder

In Germany, the key public entities regarding emergency food supply are the BMEL and the BLE. According to *DE1*, the BLE is responsible for monitoring food emergency preparedness under the jurisdiction of the BMEL. This responsibility encompasses procuring and managing national crisis reserves (DE1, p. 2, lines 29-33). These reserves are maintained through contracts with private companies responsible for storage services (DE1, p. 20, lines 8-9). However, *DE2* indicates that preparatory measures do not involve private companies from the food retail sector (DE2, p. 7, line 28). During crises, the BLE facilitates the retrieval of stored goods by the federal states. Based on *DE1*, the distribution to grain or hulling mills and the population falls under the jurisdiction of the federal states. *DE1* describes the BLE's role as a facilitator and notes that during wartime, confiscations and government orders are altering the dynamics of private ownership:

> *"In principle, we open the doors and say: 'Now here [...] we have this quantity. You can come by with trucks and pick it up'. [...] At that moment, private ownership ceases to exist. Everyone must be aware of that."*
> (DE1, p. 4, line 53 & p. 9, lines 26-29)

In Switzerland, the relevant public entities are the EAER and the FONES. *CH1* states that the FONES, associated with the EAER, oversees the execution, monitoring, and potential release of mandatory reserves maintained by private companies (CH1, p. 3, lines 22-23). Consequently, according to *CH1*, the FONES assesses whether an inventory release is warranted during a shortage scenario. This assessment includes determining whether the shortage applies to the entire industry or is limited to a specific company that may have miscalculated (CH1, p. 12, lines 15-27). After that, the FONES can release 20% of stocks in the first phase (Stage A) without Federal Council consultation, enabling quick intervention to stabilize the market (CH1, p. 12, lines 28-34). If a severe and prolonged shortage necessitates releasing more than 20% of mandatory stocks, EAER approval through an ordinance is required (CH1, p. 12, lines 16-24). Therefore, as per *CH1*, the FONES collaborates closely with approximately 60 food industry experts within the private-sector cooperative Re´servesuisse to influence legislative processes for crisis response. The FONES has delegated the calculation and allocation of storage amounts in the food sector to this cooperative (CH1, p. 18, line 44 – p. 19, line 1). The members include all companies obligated to store mandatory reserves for national economic supply. On behalf of EAER, Re´servesuisse issues general import authorizations, monitors mandatory reserve stocks, and acts as an intermediary between the private sector and authorities (CH2, p. 7, line 31-43). The cooperative manages guarantee funds provided by the federal government for financing mandatory reserve storage, compensating members for storage costs, and addressing price fluctuations (CH2, p. 11, lines 41-51). Oversight of guarantee fund use and procurement falls under the responsibility of the FONES (CH1, p. 3, lines 21-23). The concept and interrelationships of the different actors in Germany and Switzerland are shown in detail in Figure 2.

In summary, Germany and Switzerland manage emergency food supply systems through key public entities with lean administrative structures. Despite prior personnel reductions, both nations are recruiting new staff, emphasizing the

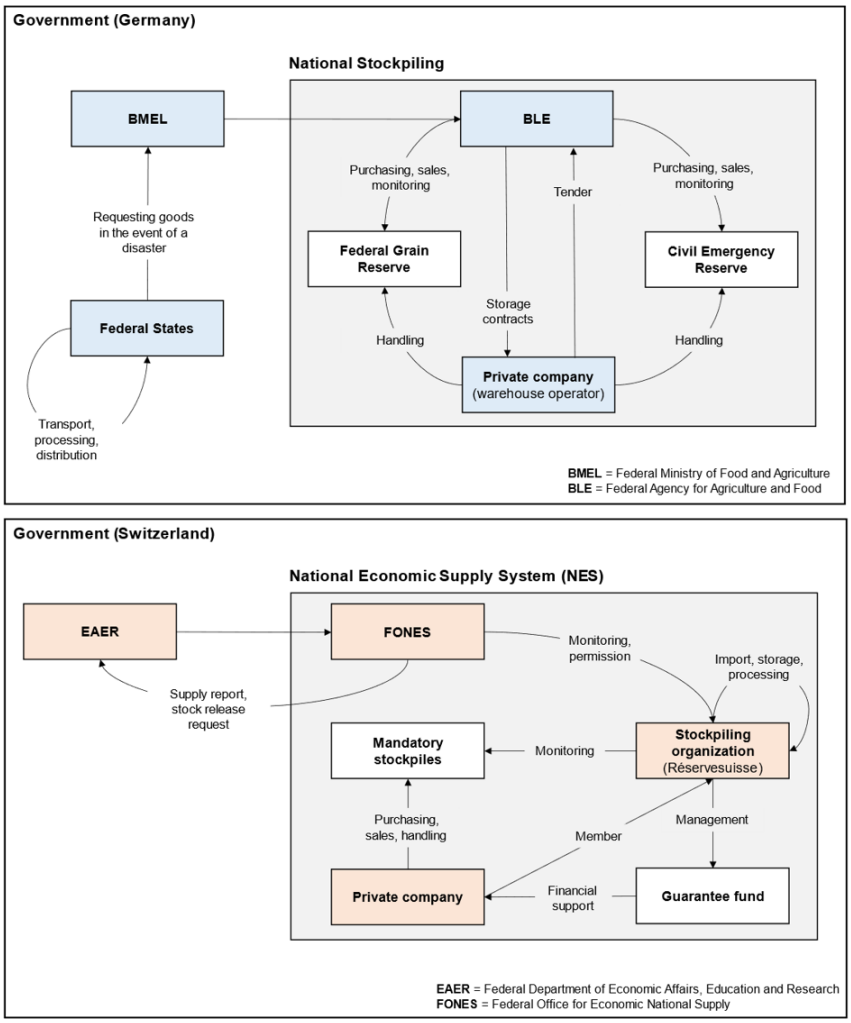


Figure 2. Organizational structure of stockpiling systems in Germany and Switzerland

increasing significance of the topic. In Germany, the BMEL and the BLE play central roles, with private companies responsible for storage services. In case of need, seizing companies' private property is also possible. There is no preventive involvement of logistics companies in this process.

On the other hand, Switzerland's EAER and FONES manage mandatory reserves maintained by private companies, mainly through the cooperative Re´servesuisse. Decisions and assessments are coordinated with representatives from the private sector, shaping mandatory storage in alignment with companies' current needs. The objective during a crisis is to maintain operations within the supply chains for as long as possible without government intervention, facilitated by the utilization of mandatory stocks.

While both countries aim to ensure a stable food supply during emergencies, the structures and mechanisms differ, reflecting the variations of governmental and private sector collaboration. Understanding these distinctions is crucial for comprehending the resilience and responsiveness of each nation's emergency food supply system.

Economic Aspects

Based on *DE1*, the federal government covers the costs of Germany's emergency food supply, focusing on crisis preparedness rather than profit generation (DE1, p. 11, lines 4-7). The storage costs, approximately =C0.33 per capita, amount to around =C27 million for Germany's population of 82 million (DE1, p. 22, lines 28-29). Funding comes from the federal budget, enabling the BLE to cover storage expenses and potential sales losses (DE1, p. 12, lines 24-26). These expenses include storing goods with private companies over an agreed period. The BLE procures and sells the stored goods in a ten-year cycle, with the expenses financed through loans (DE1, p. 10, lines 39-42).

CH1 estimates the Swiss system costs around CHF 20 million, translating to approximately CHF 6 per person per year for the country's nearly nine million inhabitants. The primary cost driver is storage expenses, including

mandatory capacities and administrative efforts, with storage costs being the exclusive expense for companies and the main factor influencing overall system costs (CH1, p. 6, lines 38-39 & p. 8, lines 35-36). As per *CH1*, only some private companies within a sector maintain mandatory stocks due to the considerable administrative burden involved. However, fair competition is ensured, as all companies must pay taxes on goods they import, and those holding compulsory stocks are reimbursed for their associated costs (CH1, p. 16, lines 13-19). *CH2* indicates that this storage method should be *"cost-neutral for mandatory reserve holders"* (CH2, p. 12, line 15). Achieving cost neutrality is crucial for encouraging companies to participate in emergency preparedness. By accessing federally guaranteed bank loans, companies can finance mandatory reserve quantities with reduced interest costs and benefit from higher tax deductions on reserve goods (CH1, p. 8, line 50 – p. 9, line 4). Moreover, *CH2* states that the funding for the system relies on robust border protection safeguarding agricultural prices. When the global market price drops below a specified threshold, customs duties are imposed at the border on imported goods, aligning the import price with the domestic price and increasing the cost of the final product. A portion of import levies, averaging CHF 40 per tonne, contributes to the guarantee fund of Re`servesuisse. This fund reimburses compulsory stockpilers for their incurred costs (CH2, p. 11, line 40- 50 & p. 12, line 6 - 7).

In economic terms, Germany and Switzerland exhibit distinct approaches to funding their emergency food supply systems. In Germany, the federal government assumes financial responsibility, allocating funds from the federal budget to cover the storage costs. A four-year budget plan is created, and the final funding amount is politically determined, influenced by global events affecting potential disposal losses. In addition, the BLE manages procurement and sales in a ten-year cycle, with expenses financed through loans. The emphasis is on crisis preparedness rather than profit generation.

In contrast, Switzerland's system costs are far surpassing Germany's costs. Storage expenses, including mandatory capacities and administrative efforts, constitute the primary cost driver. Financing involves import duties and tax benefits, focusing on cost neutrality and encouraging companies to participate in emergency preparedness. These divergent funding models reflect the unique economic strategies employed by Germany and Switzerland.

System Comparison: Perceived Benefits and Drawbacks

Germany and Switzerland implement fundamentally different approaches to ensure food supply. In Germany, the current strategy prioritizes state-supplied food within the Basic Law's framework of public welfare responsibility without specifying a definite period. In contrast, Switzerland pursues a collaborative approach between the federal government and businesses. According to the law, businesses are primarily responsible for supplying the population and maintaining mandatory stockpiles equivalent to three to four months. The two countries diverge in their current strategies, moving away from their early similarities. Recently, both countries have demonstrated an increased recognition of the significance of emergency food preparedness. Table 3 presents an overview of the comparative analysis of both systems regarding their perceived benefits and drawbacks.

The food emergency system in Germany features a robust infrastructure, ensuring affordable storage quantities to safeguard food security in disaster scenarios supported by transparent legal regulations and a streamlined administrative process. Furthermore, the system clearly defines applicability and decision-making authority while keeping storage locations confidential to mitigate the risk of looting or unfair distribution. The control over the strategy lies with the state, providing centralized oversight. In the view of *DE1*, the food reserves function as insurance for events like the COVID-19 pandemic or an event such as the Ukraine-Russia war, underscoring the rapid emergence of crises with unpredictable magnitudes (DE1, p. 22, line 14-31).

However, Germany's system has challenges. A significant drawback lies in the lack of regular adjustments and the detachment from market dynamics, unlike Switzerland, where quantities are adjusted every four years. *DE1* clarifies that in Germany, there are no regular adjustments in this regard.

> *"The most recent development in 2011 involved receiving a directive from the ministry specifying the quantities to be maintained in the reserves. This directive serves as our basis, and we must implement it annually."*
> (DE1, p. 4, lines 1-13)

Moreover, the reserves are only activated when the market ceases to function. This strategy renders using reserves for market support impractical and may complicate processing and distribution. In addition, the restricted involvement of the private sector and the minimal collaboration with other stakeholders raise questions about the effectiveness of distributing goods in an emergency and the timely availability of adequate human and transport resources. From the perspective of *DE1*, the issue arising from this system is that the BLE can *"only ever draw on the capacities that the*

Country	Category	Benefits	Drawbacks
Germany	Strategy	Food security for crises Low administrative effort Regulation of fair allocation	Lack of regular adjustments Application only in case of market failure Requirement of additional crisis processes Away from the market
	Stakeholder	Control lies with the state Clear responsibilities	Limited private sector collaboration Potential shortage of resources Federalism requires extensive coordination
	Economic	Warehousing at low cost	Strong dependence on state budget Risk of sales losses and quality problems Resource acquisition through confiscations
Switzerland	Strategy	Food security for crises Integration into the supply chain High product variety	High administrative effort Risk of looting and unequal distribution
	Stakeholder	Close collaboration Market participants offer expertise Market situation is known	Moral Hazard
	Economic	Market is protected from collapse Compensation for private sector	Strict controls and sanctions required High costs Entrepreneurial risk and competitive disadvantage

Table 3. Overview of perceived benefits and drawbacks

private sector offers" (DE1, p. 16, lines 41-42). The BLE is solely responsible for facilitating transport to potential production sites, while the actual distribution to the population falls under the jurisdiction of the federal states. In Germany, this decentralized approach presents a significant issue, as federal state jurisdiction hinders direct federal government intervention in state matters. As *DE1* points out, this federalism results in a substantial coordination requirement, which has been historically underestimated and exists only to a partial extent (DE1, p. 21, lines 41-45). Furthermore, *DE2* highlights the significance of a clearly outlined plan for distributing relief items during crises, noting that this information has yet to be communicated to the private sector. The approach entails predetermined contact points, defined responsibilities, and efficient coordination to ensure swift assistance, prioritizing essential food items and hygiene products (DE2, p. 8, lines 20-43). Aside from coordination issues, there is also the matter of whether the state possesses the necessary resources compared to private companies for efficiently managing goods and maintaining quality. While Germany benefits from low-cost warehousing, it faces challenges due to a strong dependence on the state budget and potential sales losses.

On the other hand, Switzerland adopted a decentralized approach to its food emergency system by integrating the stockpiles into the regular supply chain and offering a high product variety. It collaborates with federal and private stakeholders, allowing tailored responses to regional and unique nutrition needs like gluten intolerance. Furthermore, the close collaboration and coordination among stakeholders provide benefits, including transparency, market knowledge, logistical resources, and consideration of the interests of all stakeholders.

CH1 claims that a significant advantage lies in the strategic placement of stocks directly within the producing companies along the value chain rather than in remote locations:

> *"The warehouses are seamlessly integrated into regular production processes, not tucked away in a corner. This integration eliminates issues related to handling, such as spoilage. [...] The benefit of this approach is that all mandatory stockpiles are seamlessly integrated into the established channels. In the event of a shortage, for instance, if a particular type of bread is unavailable, the producer with compulsory stock can source from there and continue production. This process ensures that everything flows through the regular channels without needing modifications, which is a significant advantage."* (CH1, p. 4, lines 36-43 & p. 13, lines 45-51)

Consequently, questions regarding delivery destinations, responsible entities, and processing of goods become irrelevant, as the stocks are conveniently located within the companies engaged in activities like grain milling. *CH2* reinforces this claim by stating that market participants, in his perspective, clearly understand the demand for a product, including where, when, and in what quantity it is required. Moreover, they ensure delivery by established quality standards (CH2, p. 13, lines 21-25). In his view, the federal government faces the potential challenge of

accurately assessing market dynamics. It is possible to be right next to the market and not recognize it, which *CH2* identifies as the most significant risk associated with insufficient collaboration with the private sector (CH2, p. 13, lines 26-33). Furthermore, timely intervention can prevent a complete collapse and, according to *CH1*, also deter panic buying (CH1, p. 12, lines 16-20).

Nonetheless, Switzerland faces some drawbacks. The cost and administrative effort of maintaining strategic food stockpiles can be substantial, and the private nature of its system may lead to challenges like the risk of looting and unequal distribution. Moreover, while product diversity brings advantages, it also results in elevated administrative efforts and, consequently, higher expenses. As stated by *CH1*, the past was characterized by simplicity, with only a few grain variations available. However, numerous other varieties exist today, necessitating multiple silos to store each variety separately (CH1, p. 7, lines 34-52). Additionally, there is a moral hazard for companies to refrain from maintaining extra safety stocks, relying on the government to release reserves in times of scarcity. In this regard, *CH1* states that:

> *"Certain companies may have optimization considerations, questioning the need to maintain an additional month of free stock for their company. Some may argue that one week is sufficient, relying on compulsory stocks in case of emergencies. Evaluating the pain threshold in each case becomes a somewhat challenging task."*
> (CH1, p. 9, lines 46-51)

Thus, strict controls are crucial to prevent unauthorized stock releases during shortages. In 2018, as per *CH2*, his company faced a situation where low water levels on the Rhine prevented them from sourcing additional goods. Meanwhile, another company possessed sizable inventory reserves, leading the federal office to categorize it as non-crisis and refrain from releasing emergency stockpiles:

> *"We faced challenges in the past, especially when the Rhine nearly dried up, causing a prolonged halt in transportation for weeks or even months. The logistics chain was severely disrupted, and we still did not get the go-ahead because other players had larger stocks. Authorities must also establish clear criteria to determine when an economic crisis has occurred."*
> (CH2, p. 8 & 9, lines 50-52, 1-27)

According to *CH2*, the regulation lacks clarity in such cases. The Swiss system, while safeguarding the market against collapse and providing compensation to the private sector, introduces a competitive disadvantage for domestic processing. This issue arises because warehousing costs are passed on to consumers through product pricing, whereas imported goods are subjected to import levies (CH2, p. 8, lines 35-50).

DISCUSSION AND CONCLUSION

In conclusion, examining food emergency preparedness in Germany and Switzerland has highlighted significant distinctions in incorporating public and private actors, particularly concerning their stockpiling strategy, structural frameworks, stakeholder engagement, and economic approach. Germany emphasizes centralized control by the state and cost efficiency, while Switzerland focuses on collaboration and integration into the supply chain. The simplicity of Germany's approach contrasts with Switzerland's more complex and interconnected strategy involving various stakeholders. The strengths and weaknesses underscore the ongoing challenge of balancing efficiency, coordination, and adaptability in disaster management and preparedness.

Germany's system demonstrates a robust infrastructure with low administrative effort and costs. However, it raises the concern of whether ample resources and production capacity would be accessible for efficient processing and distribution during a crisis. On the contrary, Switzerland follows a decentralized approach that promotes adaptability but carries the risk of economic challenges and ambiguity regarding responsibilities. The ability to ensure equitable distribution of goods during a crisis remains to be determined.

In this comparative study, our exploration of Germany and Switzerland's diverse approaches to stockpiling critical goods and engaging with public and private stakeholders reveals valuable lessons for enhancing emergency food preparedness and addressing crisis-related challenges. While the private sector plays a crucial role in Switzerland, the interview with *DE2* sheds light on the opportunities for increased communication and awareness within the private sector in Germany. There is a clear interest in greater private sector involvement with well-defined guidelines and recognizing the benefits of integrating stored goods into the supply chain. However, *DE1* maintains that increased private sector involvement is not essential, emphasizing the state's primary role.

In Switzerland, both stakeholders are satisfied with their system but express minor requests for adjustments. These adjustments include more precise regulations during crises, fair compensation for warehouse keepers in light of rising costs, and a general preference for reduced administrative effort. Furthermore, during the interviews, when asking the experts about their "ideal vision" for food emergency preparedness, a unanimous emphasis was placed on the importance of stockpiling and maintaining national reserves. The statements highlighted the significance of close collaboration and underscored the necessity for increased investments in the underlying system and preparedness efforts. Notably, authorities in Germany and Switzerland advocate for expanding stored quantities and preparedness investments, with Switzerland proposing an extension of the potential supply period from four to six months. Furthermore, expert *CH1* suggested a hybrid approach that combines elements of both strategies, involving the private sector to use existing infrastructures and knowledge while having the government handle financing to address potential economic inequalities.

Nevertheless, it is essential to recognize the study's limitations. Most data is available in German, which could restrict the depth of our exploration. The study's scope, encompassing four interviews, might only partially capture the diverse spectrum of perspectives in both countries. Furthermore, Switzerland, less populated and smaller than Germany, operates with a unique system. Therefore, the transferability of structural elements, particularly legal aspects, might pose a challenge and necessitate further consideration in future analyses.

In addition, given the complex interconnections of critical sectors supporting food storage and supply, it is essential to consider specific crisis scenarios with greater depth, enabling the thorough identification of challenges and actionable strategies to enhance crisis management. For instance, a blackout scenario significantly affects various processes within the food supply chain, ranging from production to distribution and consumption. In Germany, logistical challenges may arise in extracting and distributing goods from storage facilities. Moreover, disruptions could occur in processing the goods at mills, bakeries, and communal facilities. Similarly, in Switzerland, it is necessary to assess how local supply chains and companies would navigate such a scenario and whether critical crisis resources like power generators are available to uphold production. Conversely, in the context of a pandemic, it is crucial to prioritize the mitigation of personnel shortages due to illness and implement suitable operational measures. Furthermore, diseases that affect digestion may influence the suitability of stored food. In other crisis scenarios like large-scale floods or earthquakes, decision-makers should focus on infrastructure functionality instead, considering potential blockages or destruction of roads and warehouses.

These diverse contexts underscore the imperative need for further detailed examination of the impact of different crisis scenarios on the systems and their performance. Moreover, investigating the integration of technology solutions or information systems for resource monitoring, enhancing communication, and fostering collaboration among various stakeholders present valuable approaches to mitigate risks and strengthen the resilience of food supply chains. Therefore, our study serves as a basis for future research and emphasizes the value of examining diverse strategic approaches and technologies in national food emergency preparedness.

REFERENCES

Arnette, A. N., & Zobel, C. W. (2019). A risk-based approach to improving disaster relief asset pre-positioning. *Production and Operations Management*, *28*(2), 457–478. https://doi.org/10.1111/poms.12934

BLE. (2023a). Kritische Infrastrukturen Landwirtschaft. https://www.ble.de/DE/Themen/Landwirtschaft/Kritische-Infrastruktur/Hintergruende/hintergruende node.html

BLE. (2023b). Staatliche Vorsorge: Lagerhaltung. Retrieved November 30, 2023, from https://www.ernaehrungsvorsorge.de/staatliche-vorsorge/lagerhaltung

BMEL. (2023). Staatliche Vorsorge. Retrieved November 30, 2023, from https://www.ernaehrungsvorsorge.de/staatliche-vorsorge

BMWK. (2024). Ö lkrisenvorsorge und -management (Bundesministerium fü r Wirtschaft und Klimaschutz, Ed.). Retrieved February 21, 2024, from https://www.bmwk.de/Redaktion/DE/Artikel/Energie/mineraloel-oelbevorratung-transport-oelreserven.html

Brem, S. (2015). Critical infrastructure protection from a national perspective. *European Journal of Risk Regulation*, *6*(2), 191–199. https://doi.org/10.1017/S1867299X00004499

BVerfG. (1977). Urteil des ersten senats vom 16. oktober 1977: 1 bvq 5/77 -, rn. 1-18, (BVerfG, Ed.). Retrieved December 5, 2023, from https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/1977/10/qs19771016_1bvq000577.html

Caballero, A. M., Teng, P. P. S., & Shrestha, M. (2015). Public stockpiling and food security (S. Rajaratnam School of International Studies, Ed.). Retrieved December 5, 2023, from https://www.files.ethz.ch/isn/191682/PB150603_Public-Stockpiling.pdf

Condeixa, L. D., Leiras, A., Oliveira, F., & de Brito, I. (2017). Disaster relief supply pre-positioning optimization: A risk analysis via shortage mitigation. *International Journal of Disaster Risk Reduction*, *25*, 238–247. https://doi.org/10.1016/j.ijdrr.2017.09.007

Edge, S., & Meyer, S. B. (2019). Pursuing dignified food security through novel collaborative governance initiatives: Perceived benefits, tensions and lessons learned. *Social science & medicine (1982)*, *232*, 77–85. https://doi.org/10.1016/j.socscimed.2019.04.038

FONES. (2011). Geschichte der wirtschaftlichen landesversorgung: Seit gru¨ndung des bundesstaates. Retrieved July 26, 2023, from https://www.bwl.admin.ch/bwl/de/home/wirtschaftliche_landesversorgung/prasentation_wl/geschichte.html

FONES. (2018). Strategische Ausrichtung Strategische Ausrichtung der wirtschaftlichen Landesversorgung. Retrieved December 5, 2023, from https://www.bwl.admin.ch/dam/bwl/de/dokumente/dokumentation/publikationen/strategie 2018.pdf.download.pdf/Strategie WL D 2018.pdf

FONES. (2021). Bericht zur wirtschaftlichen Landesversorgung 2017-2020. Retrieved July 26, 2023, from https://www.newsd.admin.ch/newsd/message/attachments/66655.pdf

FONES. (2022). Pflichtlager. Retrieved July 26, 2023, from https://www.bwl.admin.ch/bwl/de/home/bereiche/pflichtlager.html

FONES. (2023). Bericht zur Vorratshaltung 2023. Retrieved November 30, 2023, from https://www.bwl.admin.ch/bwl/de/home/bereiche/pflichtlager.html

Gerhold, L., Cortez Garcia, K., & Guerrero Lara, A. (2019). *Lebensmittelversorgung in Krisen und Katastrophen: Versorgung der Bevo¨lkerung mit Lebensmitteln in OECD-La¨ndern im Falle von Großschadensereignissen* (Vol. Band 22). Bundesamt fu¨r Bevo¨lkerungsschutz und Katastrophenhilfe.

Kumar, M., Raut, R. D., Sharma, M., Choubey, V. K., & Paul, S. K. (2022). Enablers for resilience and pandemic preparedness in food supply chain. *Operations Management Research*, *15*(3), 1198–1223. https://doi.org/10.1007/s12063-022-00272-w

Manzoor, U., Zubair, M., Batool, K., & Zafar, B. (2014). A multi-agent framework for efficient food distribution in disaster areas. *International Journal of Internet Technology and Secured Transactions*, *5*(4), 327. https://doi.org/10.1504/IJITST.2014.068711

Mayring, P., & Gahleitner, S. (2010). Qualitative inhaltsanalyse. https://doi.org/10.2307/j.ctvhktjdr.31

Menski, U. (Ed.). (2016). *Neue Strategien der Erna¨hrungsnotfallvorsorge: Ergebnisse aus dem Forschungsverbund NeuENV* (Vol. Nr. 18). Freie Universita¨t Berlin, Forschungsforum Öffentliche Sicherheit.

Mor, R., Sid, S., Panghal, A., Kumar, D., & Kumar, V. (2021). Agri-food supply chain and disruptions due to covid-19: Effects and strategies. *Brazilian Journal of Operations & Production Management*, *18*, 1–14. https://doi.org/10.14488/BJOPM.2021.031

Ni, W., Shu, J., & Song, M. (2018). Location and emergency inventory pre-positioning for disaster response operations: Min-max robust model and a case study of yushu earthquake. *Production and Operations Management*, *27*(1), 160–183. https://doi.org/10.1111/poms.12789

Ö zdamar, L., Ekinci, E., & K ü ç ü kyazici, B. (2004). Emergency logistics planning in natural disasters. *Annals of Operations Research*, *129*(1-4), 217–245. https://doi.org/10.1023/B:ANOR.0000030690.27939.39

Rawls, C. G., & Turnquist, M. A. (2010). Pre-positioning of emergency supplies for disaster response. *Transportation Research Part B: Methodological*, *44*(4), 521–534. https://doi.org/10.1016/j.trb.2009.08.003

Rejeb, A., Keogh, J. G., & Rejeb, K. (2022). Big data in the food supply chain: A literature review. *Journal of Data, Information and Management*, *4*(1), 33–47. https://doi.org/10.1007/s42488-021-00064-0

Re´servesuisse. (2023). Re´servesuisse genossenschaft (Re´servesuisse, Ed.). Retrieved November 30, 2023, from https://www.reservesuisse.ch/reservesuisse-genossenschaft/

Rickenbacher, D. (2021). Die Schweiz in der Suez-Krise (Schweizerisches Nationalmuseum, Ed.). Retrieved July 26, 2023, from https://blog.nationalmuseum.ch/2021/11/die-schweiz-in-der-suez-krise/

Sabbaghtorkan, M., Batta, R., & He, Q. (2020). Prepositioning of assets and supplies in disaster operations management: Review and research gap identification. *European Journal of Operational Research*, *284*(1), 1–19. https://doi.org/10.1016/j.ejor.2019.06.029

SE. (1982). Bundesgesetz vom 8. Oktober 1982 u¨ber die wirtschaftliche Landesversorgung (Landesversorgungsgesetz, LVG. Retrieved July 26, 2023, from https://www.fedlex.admin.ch/eli/cc/1983/931 931 931/de

SE. (2023). Landesversorgungsgesetz-Teilrevision wird Erfahrungen aus der Energiekrise einbeziehen. Retrieved July 26, 2023, from https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-92500.html#:~:text=Der%20Bundesrat%20beschloss%20am%2030,neuen%20und%20gestiegenen%20Anforderungen%20anzupassen.

Shrivastav, S. K., & Bag, S. (2023). Humanitarian supply chain management in the digital age: A hybrid review using published literature and social media data. *Benchmarking: An International Journal*. https://doi.org/10.1108/BIJ-04-2023-0273

SWI. (2022). Wie ging die Schweiz mit der Ö lkrise 1973 um? Retrieved July 26, 2023, from https://www.swissinfo.ch/ger/kultur/energiekrise_ wie-ging-die-schweiz-mit-der-oelkrise-1973-um--/47920866

Verma, A., & Gaukler, G. M. (2015). Pre-positioning disaster response facilities at safe locations: An evaluation of deterministic and stochastic modeling approaches. *Computers & Operations Research*, *62*, 197–209. https://doi.org/10.1016/j.cor.2014.10.006

Weersink, A., von Massow, M., Bannon, N., Ifft, J., Maples, J., McEwan, K., McKendree, M. G., Nicholson, C., Novakovic, A., Rangarajan, A., Richards, T., Rickard, B., Rude, J., Schipanski, M., Schnitkey, G., Schulz, L., Schuurman, D., Schwartzkopf-Genswein, K., Stephenson, M., . . . Wood, K. (2021). Covid-19 and the agri-food system in the united states and canada. *Agricultural Systems*, *188*, 103039. https://doi.org/10.1016/j.agsy.2020.103039

WEF. (2023). The global risks report 2023. Retrieved June 1, 2023, from https://www.weforum.org/reports/global-risks-report-2023/

Willisegger, J. (2015). Liberalismus oder Staatsintervention. Die Geschichte der Versorgungspolitik im Schweizer BundesstaatCottier, MauriceZu¨ rich, Verlag Neue Zürcher Zeitung (2014), 216 S., ISBN 978–3–03823–900–0. *Swiss Political Science Review*, *21*(3), 470–474. https://doi.org/10.1111/spsr.12173

Empirical Investigation of Values Affected by AI Systems for Fire Services

Tina Mioch*
Delft University of Technology
University of Applied Sciences Utrecht
tina.mioch@hu.nl

Huib Aldewereld
University of Applied Sciences Utrecht
huib.aldewereld@hu.nl

Mark A. Neerincx
Delft University of Technology
M.A.Neerincx@tudelft.nl

ABSTRACT

Fire fighters operate in a dangerous, dynamic, and complex environment. Artificial Intelligence (AI) systems can contribute to improve fire fighters' situation awareness and decision-making. However, the introduction of AI systems needs to be done responsibly, taking (human) values into account, especially as the situation in which fire fighters operate is uncertain and decisions have a big impact. In this research, we investigate values that are affected by the introduction of AI systems for fire services by conducting several semi-structured focus group sessions with (operational) fire service personnel. The focus group outcomes are qualitatively analyzed and key values are identified and discussed. This research is a first step in an iterative process towards a generic framework of ethical aspects for the introduction of AI systems in first response, which will give insight into the relevant ethical aspects to take into account when developing AI systems for first responders.

Keywords

Values, Fire services, Value sensitive design, Responsible AI.

INTRODUCTION

First responders (FRs) operate in a dangerous, dynamic, and complex environment, in which they have to quickly understand the situation and how to mitigate danger, while keeping themselves and civilians safe. To improve effectiveness and safety, FR needs and corresponding capability gaps were identified by the International Forum to Advance First Responder Innovations (IFAFRI), e.g., the ability to create actionable intelligence based on data and information from multiple sources (Capability Gap 9) and the ability to conduct on-scene operations remotely without endangering responders (Capability Gap 7) (IFAFRI, 2019). Technology could enhance the safety, effectiveness, and efficiency of FRs and help closing these capability gaps.

One of the technologies which has a high potential for contributing to filling these capability gaps is artificial intelligence (AI). AI systems can play an important role in the future to, for example, improve shared situation awareness (e.g., (Mioch et al., 2021)) and contribute to advanced decision-making (Radianti et al., 2019). This leaves FR organizations with the challenge of integrating these AI systems in the decision-making processes in a responsible way by addressing (public) values and public function properly. In our view, AI systems in this domain are always part of a hybrid human-AI system, a socio-technical system, in which task allocation and task responsibility might change and new human-AI dependencies arise. This introduces the research challenge of developing hybrid human-AI systems in which AI technology and humans cooperate in a way that synergy is created (Akata et al., 2020; Seeber et al., 2020). In such a hybrid intelligence system, complex goals can be accomplished by combining

*corresponding author

humans and AI technology that collectively work on shared objectives with complementary capabilities that, when combined, augment each other (Dellermann et al., 2019).

The development and application of these hybrid human-AI systems need to be done responsibly (AI HLEG, 2018), e.g., regarding possible biases in (training) data sets and privacy aspects (e.g., (personal) data on FRs such as location, performance, stress). For some domains, the ethical aspects of AI systems and applications have received a lot of attention, e.g., the health domain (Blasimme & Vayena, 2020; Morley et al., 2020; Murphy et al., 2021) and the military domain (Galliott & Scholz, 2020; Wasilow & Thorpe, 2019). However, in the field of FR, it seems that ethics with regard to the application of AI has not yet been much addressed. To develop and apply AI systems responsibly, it is important that the AI systems support the stakeholders' values and that values of different relevant stakeholders are taken into account. The context in which an AI system is applied and choices in the (technical) development of the AI system determine the relevance and importance of the different values. In previous work in the FR domain, key ethical concerns have been identified for Search and Rescue robots to support development and deployment in a responsible way (e.g., see Harbers et al. (2017) for empirical research and Battistuzzi et al. (2021) for a scoping review). Following on this work, in this paper, we provide a qualitative empirical investigation of relevant values of relevant stakeholders of AI systems for fire services. This is a first step towards a generic framework of ethical aspects for AI in FR, which will give insight into the relevant ethical aspects to take into account when developing AI systems for FR. Subsequent steps for setting up the generic framework are differentiating between general values and instances of these values in specific use cases to take the specific context into account and linking and updating the framework with societal developments regarding ethical aspects such as norms and guidelines.

This research is inspired by the Value Sensitive Design (VSD) methodology, which accounts for human values throughout the design process (Friedman & Hendry, 2019). First, we identify relevant direct and indirect stakeholders of fire services. Second, we investigate relevant values of these stakeholders regarding the application of AI systems. To do this, we conducted three focus group sessions with fire service personnel in which we assess and analyze the stakeholders and their values. Using the focus group session results, we identify a first set of key values to take into account in the application of AI systems for fire services.

In the following, we first give background on ethical themes for AI systems and on values and value-sensitive design. We then describe the setup and execution of our focus group sessions and the results of these sessions, followed by the conclusion and discussion of the results.

BACKGROUND

In the last few years, a lot of attention has gone towards the responsible development of AI, amongst others by the EU High-level expert group on AI (AI HLEG, 2018)) and IEEE (IEEE, 2021). The AI HLEG was appointed to advise on a European AI strategy and identified key requirements, which each AI technology needs to fulfil before it is considered safe and trustworthy (AI HLEG, 2018). According to these guidelines, trustworthy AI systems should be lawful (i.e., respecting all applicable laws and regulations), ethical (i.e., respecting ethical principles and values), and robust (i.e., both from a technical perspective while taking into account its social environment). Four ethical principles are identified that AI systems should adhere to, namely respect for human autonomy, prevention of harm, fairness, explicability, and 7 key requirements, i.e., human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental wellbeing, and accountability. These requirements are explicitly non-exhaustive. When applying these guidelines to specific AI applications, the ethical principles need to (1) be situated and considered in context to determine their relevance and supplement them with possible additional ethical requirements and (2) be translated into design requirements for the specific application and context. The AI HLEG does not give guidelines on how to do this and they have not been applied structurally to the FR domain.

Technology is not neutral but always manifests values (van de Poel, 2015). Values refer to what a person or group of people considers important in life (Friedman et al., 2013). Schwartz (2012) identified 10 basic universal values, i.e., power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity, and security. Some of these values conflict with each other (e.g., benevolence and power), whereas some are compatible with each other (e.g., conformity and security). A possible approach to support taking values into account in technology design is Value sensitive design (VSD). VSD is "a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process" (Friedman et al., 2013, p. 56). Values that are considered particularly important for technological design are Welfare, Ownership and Property, Privacy, Freedom from Bias, Universal Usability, Trust, Autonomy, Informed Consent, Accountability, Courtesy, Identity, Calmness and Environmental Sustainability. As we are

investigating values affected by AI systems for fire services and not general human values of stakeholders within the fire services, we use Friedman et al.'s account of values as a starting point in our investigation.

VSD is an iterative methodology that integrates three perspectives, namely conceptual, empirical, and technical investigations. These investigations are executed iteratively, though not necessarily in a set order. VSD contains a rich collection of different methods that help designers to investigate values in technology, such as direct and indirect stakeholder analysis and value scenarios (Friedman et al., 2017), and is thus particularly useful for our research. To be able to determine human values that should be taken into account, VSD asks system designers to establish a robust set of stakeholder groups and to justify those who likely are most strongly affected – that is, to provide an conceptual or empirical rationale for their inclusion in the design process. The same holds for values; the applicability and relevance of values should be investigated from the three different perspectives. In this research, we focus on one of VSD's perspectives, namely empirical investigation, by qualitatively identifying a set of stakeholders to include and key values to take into account when developing AI for fire services. In later steps, we will iteratively extend, refine, and test the results, also from a conceptual and technical perspective.

A tool that can help with creating an overview of the potential positive and negative impact of AI systems is the *Ethical matrix* (Mepham et al., 2003). The Ethical matrix is a conceptual tool designed to help decision makers deliberate about the ethical acceptability of existing or envisioned technology. It makes the impact of design choices on the different stakeholders explicit and provides structure and support in the design process. The cells of the matrix contain the impact, negative or positive, of the envisioned technology, on each of the stakeholder groups for specific ethical values.

METHOD

To determine relevant values for fire services regarding the application of AI, we conducted semi-structured interviews during focus group sessions with different subject-matter experts from the fire services. The goal of the focus group sessions was exploratory, to gain as much information as possible on expected impact of AI systems for fire services, and the results were qualitatively analysed.

The first focus group was conducted as an exploratory session with three incident commanders that are involved in innovation projects. The goal was twofold, first, to start exploring relevant stakeholders and expected impact of AI systems on these stakeholders, second, to test and evaluate the focus group setup. Two other focus group sessions with each 4 participants were conducted with fire fighters (carrying out day-to-day firefighting and fire safety work; 3 participants), dispatchers (managing emergency calls, ensuring that proper response teams are sent to the incident location; 2 participants), incident commanders (in charge of an incident; 2 participants), and an HR professional (1 participant). All participants (besides the HR professional) had experience in the field, ranging from limited experience (1-3 years experience) to very experienced (15+ years of experience). The focus group sessions lasted 2 hours.

The first, exploratory, focus group session consisted of several steps. To identify as many stakeholders as possible, the participants were asked to list all people that interact with an envisioned AI system (presented in a general scenario) and people that are affected by the system or who have a vested interest in its success or failure. To not have any influence effects on the process, the experts were asked to do this by themselves. After 5 minutes, the results were shared and discussed. Based on the combined list of stakeholders, a prioritization of stakeholders was made. After having selected the most relevant stakeholder, the participants determined which positive or negative impact of AI systems they could identify for these stakeholders. The participants went through the stakeholders (by themselves) and wrote identified impact on post-its. The results were plenary shared and discussed; as every participant shared their identified impact, participation of each of them was ensured. Additional impact was added during the discussion. After the focus group session, the identified (positive and negative) impact was analyzed regarding (possible) underlying values by a human-AI collaboration expert, mapped on an ethical matrix (van der Stappen & Steenbergen, 2020) and presented to the participants for validation. The discussion of this first version of the ethical matrix led to additional input regarding impact for stakeholders.

The other two focus group sessions consisted of the same steps as the first exploratory focus group, with slight adaptations, based on lessons-learned of the first focus group session: we realized that (1) a general knowledge of the working of AI systems is needed to identify impact of the systems and (2) example scenarios support the identification of impact. The participants of the first focus group were all involved in (AI) innovation projects and familiar with the possibilities of AI technology; for the other two focus groups, this was not the case. For that reason, we added a short overview over AI technology, its working, possibilities, and limitations. Also, two concrete scenarios (inspired by examples mentioned during the exploratory focus group session) were introduced as basis for the discussion on positive and negative impact of AI systems. These (descriptive) scenarios highlighted value

Table 1. Scenarios used during the second and third focus group.

	Scenario 1	*Scenario 2*
Situation	Large fire in storage tank; Fire brigade team is exploring the situation. Fire fighters wear sensors.	Tank truck with hazardous material collides with a tree within city limits. Hazardous substances spill.
Data	Measurement of physical properties of firefighters such as stress level, position, body temperature, presence of dangerous gases, body cam.	Video data from road cameras is processed automatically. Data from past incidents is available.
AI technology (first scenario version)	An AI system processes the data in real-time and combines the various data sources; it provides information to the incident commander about the current status of the firefighters in the field, monitors the situation and stress, and warns the incident commander in case of conspicuous things, e.g., if fire fighters come too close to a danger zone.	An AI system reports the incident to the dispatcher; it recognizes hazardous substances through ADR sign recognition and passes on this information. It also predicts the development of the situation based on similar incidents in the past and recommends to the dispatcher which resources should be sent. The dispatcher forwards reports and information to the fire house.
AI technology (second scenario version, extended autonomy of AI system)	Continuously monitors and predicts the (stress) status of the firefighters, also by means of stress-related data from the past. System advises the incident commander on employability (short-term and long-term)	The AI system reports the incident to the dispatcher, assesses who needs to be called, and puts the call through to the responsible fire house. It recognizes hazardous substances through ADR sign recognition and also passes on this information. It predicts situation development based on similar incidents in the past and forwards call for appropriate resources. The dispatcher monitors the report.

tensions. We described our scenarios as 'what if' scenarios, scenarios that are situated just a few years into the future and inviting the participants to assess ethical issues posed in the scenario (Wright et al., 2014), see Table 1. Each scenario consisted of two versions, with the second version extending the AI's autonomy to increase (possible) ethical issues.

The identified impact of AI systems was analyzed through thematic analysis and mapped onto affected underlying values. We coded inductively as a way to enter the data analysis with a more complete, unbiased look at the themes throughout the data. We categorized the resulting 28 codes into 13 themes, which will be described in the next section.

RESULTS

In this section, we describe the results of the focus group sessions. First, an overview is given of the most important stakeholder groups that were identified by the participants, together with the human values (as identified by Friedman et al. (2013) that are affected by the identified impact. Subsequently, the identified impact on the human values is described in more detail and in context of fire services, followed by other, more AI-related values that participants mentioned regarding the introduction of AI systems for fire services.

Stakeholders and Values

During the focus groups, participants identified the following stakeholders as most relevant to take into account during the introduction of AI systems: fire fighters, incident commanders, special operations fire fighters (e.g., for hazardous materials and digital exploration), dispatchers, company doctors, incident researchers, citizens, and the surroundings. For these stakeholders, the participants identified possible positive and negative impact of AI applications. Most impact was identified for three stakeholder groups, namely fire fighters, incident commanders, and dispatchers. This is not surprising, as the scenarios involved these stakeholder groups explicitly, and these stakeholder groups were (mostly) represented by the participants during the focus groups. Table 2 provides an overview of the (most relevant) stakeholders as identified during the focus group sessions together with the human values (as specified by Friedman et al. (2013)) which the identified impact affects.

Table 2. Most relevant stakeholders as identified during the focus group sessions and values (as specified by Friedman et al. (2013)) that the identified impact affects.

Stakeholders	*Values*
Incident commanders	Autonomy, identity, physical well-being
Fire fighters	Autonomy, identity, informed consent, privacy, physical well-being, psychological well-being, trust
Dispatchers	Autonomy, identity, privacy, physical well-being, psychological well-being
Citizens	Privacy, physical well-being, psychological well-being, trust
Surroundings	Physical well-being
Special operations fire fighters	Physical well-being
Incident researchers	Physical well-being
Company doctors	Physical well-being, psychological well-being

Human values

In this section, we describe in more detail which impact participants mentioned and the different human values that are implicated, based on the human values as described and identified by Friedman et al. (2013). For an overview of the impact mapped on stakeholders and affected human values, see Figure 1.

Autonomy Autonomy refers to people's ability to decide, plan, and act in ways that they believe will help them to achieve their goals (Friedman et al., 2013). Regarding autonomy, in general, participants were concerned over their dependence on technology. If FRs become too dependent on the technology, their ability to operate will be substantially impaired when systems fail. In addition, a concern is staying in control of decisions.

Regarding dispatchers, three aspects were mentioned. AI might become more and more autonomous and take over several tasks of the dispatcher. This could (1) lead to less autonomy in their task execution, and (2) impact the decision-making power of the dispatcher, and (3) disproportional trust of the results or advise of AI systems, with dispatchers reflecting less on the situation and thus operating less autonomously.

Regarding fire fighters, it is expected that AI systems lead to new needs regarding capabilities and knowledge. This might (negatively) impact their ability to act and make decisions in their operations. Furthermore, AI systems will also have impact on their task decision-making power, depending on the level of autonomy of AI systems regarding decision-making. In addition, because possible access to historic data, AI systems can give advise on (long- and short-term) employability, which might impact fire fighters' control about own employability, with possibly others deciding on their performance and health. Participants are concerned that there will be less space to make decisions based on own insights, and that human aspects, context, and experience will be less taken into account when making decisions.

For incident commanders, participants mention that AI could lead to the feeling of not being in control of own decisions anymore. Participants were also worried regarding the influence of AI advise on the decisions of incident commanders. Also, they mentioned that the human aspects remain very relevant in decision-making; for example, optimal teams based on some objective measure are not always the best, as other goals of teams should be taken into account, such as that people can learn and grow.

No positive impact of AI systems on autonomy was mentioned.

Identity Identity refers to people's understanding of who they are over time, embracing both continuity and discontinuity over time (Friedman et al., 2013). Participants mentioned that AI systems will have impact on the capabilities, knowledge, and possibly level of education that is needed for the job (of incident commanders, fire fighters as well as dispatchers). For example, AI systems might become more and more autonomous and might take over decisions that at the moment are taken by dispatchers. The participants saw this as a threat for their understanding of their work. For fire fighters and incident commanders, participants also mentioned that a culture change is needed, as AI gives completely new possibilities, which means that operations will need to change (e.g., regarding gaining situation awareness first remotely by means of autonomous UGVs and UAVs, instead of through human operation). In addition, participants expect that new (commander) roles will be created, e.g., a role as specialist digital exploration.

	Autonomy	Identity	Informed consent	Privacy	Physical well-being	Psychological well-being	Trust
Incident commander	- Dependence on AI - Decision-making - Human aspects	- Culture change is needed			+ Can monitor safety of team + Better long-term employability + Safer incident response		
Fire fighter	- Ability to make decisions - Control on employability - Human aspects and experience - Dependence on AI	- Needed capabilities, knowledge, and level of education will change - Culture change is needed	- Unforeseen use of data	- Unforeseen use of data - Who has access to data? - Feeling of being closely monitored	+ Better insight in own health status + Less stress + Better health + Better long-term employability + Safer incident response - (Felt) unpredictability of UGVs - Physical impairment	+ Less stress - Less social environment	- Feeling of being closely monitored
Dispatcher	- Task execution - Decision making - Reflection	- Needed capabilities, knowledge, and level of education will change - Threat for work		- More access to sensitive data	+ Better insight in own health status + Safer incident response through faster alarming	+ Less PTSD + Less stress	
Citizen				- Personal data	+ More safety through better recognition of environmental effects	- Less personal contact	- Feeling of being monitored
Surroundings					+ More safety through better recognition of environmental effects		
Special ops					+ More safety through better insight in presence dangerous substances		
Incident researcher					+ More safety through better insight and analysis on prevention		
Company doctor					+ Better overview of health status + Better early detection	+ Better overview of stress and trauma	

Figure 1. Ethical matrix. Red cells denote negative impact on the corresponding value, green cells positive impact, and yellow cells positive as well as negative impact.

Informed consent Informed consent refers to "garnering people's agreement, encompassing criteria of disclosure and comprehension (for 'informed') and voluntariness, competence, and agreement (for 'consent')" (Friedman et al., 2013). The participants mentioned that the introduction of AI systems could have impact on informed consent. It might not be clear what exactly is being done with (personal) data that is collected of fire fighters and that, when available, this data could be used for purposes that were not foreseen. The participants mentioned that a possible solution is the introduction of personal data safes which are not generally accessible and that people can determine in which situations these data safes are accessible and to whom (e.g., in life-critical situations to their incident commanders).

Privacy Privacy refers to a "claim, an entitlement, or a right of an individual to determine what information about himself or herself can be communicated to others" (Friedman et al., 2013). Impact on privacy was mentioned for fire fighter, dispatchers as well as citizens. For fire fighters, impact on privacy is seen mostly regarding two aspects, i.e., that of proportionality and of surveillance. First, when collecting and analyzing (personal) data, it is not always clear what the goal is of doing so and what the data is actually used for. There must be clear agreements on what AI systems do with the data and who has access to it (and to the results of a possible analysis). In addition, participants mentioned that fire fighters in general resist having their data collected, as it gives them the feeling of being closely monitored.

Dispatchers will need to take privacy (even more) into account. They will have more access to, for example, cameras on roads during incidents, and can directly watch the situation to support their situation awareness. Also, they are in direct contact with other first response organizations, such as police, that have access (and share) information during incidents.

Regarding citizens, participants mentioned that with the introduction of AI systems, care needs to be taken with personal information of citizens, such as video pictures and gps data.

Physical well-being Physical well-being falls under the value *Welfare* as identified by Friedman et al. (2013); in this context, we understand it as physical health and safety. For all identified stakeholders, participants mentioned that AI systems could impact the value *physical well-being*. AI systems could have a positive impact on the physical safety and well-being of incident commanders, fire fighters, citizens, and the surroundings as it could lead to a better situation awareness because of additional relevant information (on the incident and on personnel). This information leads to a more effective task performance through faster, earlier, more stable, and more informed decision-making of all professional stakeholders. For fire fighters, dispatchers, and incident commanders, having more information on one's own situation and being able to monitor and analyse one's physical status (e.g., regarding stress) and

acting on the status could make the job healthier and lead to better long-term employability. This holds also for the company doctor, in addition to possible early detection of (work-related) illness. For special ops and incident researchers, AI systems might lead to better insight and analysis of dangerous effects or situations, which in turn might lead to a safer incident response for all stakeholders. Also, deploying drones increases the feeling of safety.

However, physical well-being could also be impacted negatively. For fire fighters, carrying sensors or AI systems might impair movement and mobility, leading to less safe performance. In addition, being with UGVs in the same environment can increase a feeling of unsafety.

Psychological well-being Psychological well-being also falls under the value *Welfare* as identified by Friedman et al. (2013). According to the participants, for dispatchers as well as for fire fighters, AI systems can lead to improved psychological well-being. The participants mentioned that stress and PTSD (for dispatchers) could occur less because of more effective operations and because of being able to monitor one's (stress) situation.

As possible negative impacts, participants mentioned that the fire service organisation is a very social environment. A concern is that this will change with the introduction of AI systems, as human aspects might become less important, which might lead to less (psychological) well-being of fire fighters. In addition, AI systems might lead to less contact between citizens and dispatchers (due to for example automatic alarming); however, personal contact is important for citizens, to be heard and supported in a stressful situation.

Trust Trust refers to expectations that exist between people who can experience good will, extend good will toward others, feel vulnerable, and experience betrayal (Friedman et al., 2013). The human value trust is mentioned by participants to be impacted by AI for two stakeholder groups, namely fire fighters and citizens, with respect to trust in the first response organization.

For fire fighters, to effectively use AI, more data will be collected, also from fire fighters. Participants mentioned that fire fighters generally distrust the collection of (personal) data and feel closely monitored. They worry that the data is used against them (e.g., proving that a performance is not up to par). In addition, no context is present in the data, which is important to explain, for example, deviation from protocol. For citizens, trust in first response can be decreased for the same reasons. For example, the position of emergency callers is available for dispatchers, as well as of other people at the incident location. Monitoring and following these citizens (through AI systems) might enable a more effective operation; however, it could decrease trust in first response organisations due to a feeling of being monitored.

AI-related values

The participants of the focus group sessions mentioned several other values regarding the introduction of AI for fire services that are not part of the human values that are often implicated by ICT systems as identified and described by Friedman et al. (2013). This is not surprising, as Friedman's list of values relates to ICT systems in general and not AI systems in particular. The other values mentioned are *Accountability*, *Reliability*, *Trust in AI systems*, *Security*, *Transparency*, and *Appropriate training*. These values were mentioned in the context of what AI systems should adhere to, or how AI systems should function and were mostly mentioned independent of stakeholders, but as general aspects that should be taken into account during the design, development, and deployment of AI.

Accountability Accountability is seen in the context of liability (legal responsibility) as well as social responsibility. Participants are concerned that the legal responsibility is not clearly specified; aspects as liability when (wrong) decisions are taken (by the AI system as well as by the FR when being advised differently by an AI system) need to be taken into account when introducing AI systems. Also, for FRs, it is very important to be accountable for the decisions taken during operations, e.g., to be able to explain decisions after the operation. This also needs to be the case when AI systems are integrated into the operation.

Reliability Participants mentioned that it is important that AI systems are reliable, both in the results they present (e.g., in all situations, they should present working results) as well as in availability of the systems (e.g., they should not fail in high-stress or dangerous situations). This corresponds to the key requirement *Technical robustness and safety* of the HLEG (AI HLEG, 2018).

Trust in AI systems Participants mentioned that (appropriate) trust in AI systems is important. Topics that were mentioned regarding trust in AI systems were (1) understanding of system as the basis for trust, (2) *appropriate* trust, (3) dependence on systems. Regarding understanding of the system, participants mentioned that it is important to understand what the system is doing (e.g., predicting the next actions of a UGV), how the AI got to a result or advise, and how certain the advise is. Regarding appropriate trust, participants said that there is a danger of too much trust (e.g., being influenced in decision-making too much, not evaluating and reflecting enough on the advise of the AI) or of too little trust (and not taking the advise of the AI sufficiently into account). Participants were also concerned that if the AI system has some kind of autonomy (e.g., when recognizing incidents and sending relevant information about the incident when alarming first response), the FRs have no choice other than to trust the system, and thus depend on the system, as there is no time to validate all information before acting.

The participants also stressed that an incremental introduction of AI is necessary to learn about the capabilities of the AI system and how to integrate it into work processes and build up trust. This also is important because often, AI systems in practise do not work as well as in theory, which leads to distrust and less support.

Security Regarding security, participants mentioned that it is very important that data is kept at a secure place and that access control is in place. Misuse of data needs to be avoided; cybercrime is a danger.

Transparency Regarding transparency, participants mentioned several important aspects. First of all, it is important to have an understanding of the results of the AI system (e.g., advice) and the reasons why the AI system comes to these results. In addition, it is important to have an idea on the certainty of the results. Also, explanations of the advise are important, ideally taking the context into account (e.g., which context factors improve the certainty). Regarding UGVs, participants mentioned that it is important to have insight into the plan/goal of the robot, to be able to predict behaviour.

Appropriate training Regarding appropriate training, participants mentioned two topics: (1) it is of utmost importance that FRs learn how to work with AI systems, as most do not currently have sufficient IT knowledge, and (2) that it is important that FRs explicitly keep training on how to work effectively without AI systems, so that they are able to operate if the system fails and are able to evaluate advise from AI systems regarding applicability.

IMPLICATIONS ON DESIGN OF AI SYSTEMS FOR FR

In this section, we describe the implications of the results as described in the previous section on the design of responsible AI for fire services. As shown in Figure 1, positive as well as negative impact has been identified during the focus group sessions for the different human values. According to the participants, several human values are impacted (mostly) positively for the different stakeholders (although there is no human value that is only affected positively), i.e., *physical well-being* and *psychological well-being*. Several human values are impacted only negatively for the different stakeholders, i.e., *autonomy*, *identity*, *informed consent*, *privacy*, and *trust*. Particular attention should go to these values in the design-, development-, and deployment process of AI systems; AI systems should (explicitly) support the possible positive impact that they could have whereas the negative impact should either be reduced or avoided altogether, e.g., by specifying design requirements that explicitly take these human values into account. To do this, the AI-related values that have been identified can help; these were values that the participants said that AI systems should adhere to. For example, having design requirements regarding the values *accountability*, *transparency*, and *appropriate training* could help to reduce the expected negative impact on *autonomy*. Regarding *accountability*, understanding the responsibility distribution between an AI system and human would help a FR's ability to act and make decisions; regarding *transparency*, understanding advise of an AI system in the current context and being able to interpret its reliability lessens dependence on the results of the AI system; the same holds for being *trained* in how to work with and without the AI system. The translation from human values to design requirements is not straightforward (Aldewereld & Mioch, 2021). However, the values form a good basis to take into account in the design and development process of specific AI applications. For each specific AI application, the values should be re-evaluated for applicability and, together with stakeholders, value-based design requirements should be specified that promote positive impact and limit negative impact on these values. For our current research, design requirements can only be very abstract as we have not identified a specific AI system that is the subject of this research, but AI systems in general. As a proof of concept, we give an example of a translation of the value *Autonomy* into possible general design requirements, see Table 3.

Table 3. Examples of design requirements for a selection of the identified impact of AI systems on the value *Autonomy*.

Impact of AI system on Autonomy	*Design Requirements*
Dependence on AI	Personnel will[1] be extensively trained on how the AI system works and how to interact with the system during operations Personnel will be extensively trained on how to work. effectively without the AI system in case of failure and to be able to evaluate advise from the AI system. The AI system will be able to explain why a particular result or advise is given. This explanation will take the context into account and be adapted towards the particular task and role of the FR.
Decision-making	The human will stay in control of (operational) decisions. The role of the AI system in the task execution will be clear and integrated into procedures. Accountability and responsibility will be explicitly discussed and set for each AI system that is introduced.
Human aspects and experience	Protocols will be specified carefully regarding decision-making with AI systems, to make sure that human aspects and experience can be taken into account. The AI system will be able to learn from human experts.

[1] We use 'will' in the design requirements specification instead of 'shall' or 'should', as the latter imply a normative load.

CONCLUSION AND DISCUSSION

In this paper, we have made a first step towards a generic framework of ethical aspects of AI systems for FR. We held three focus group sessions with different stakeholders, discussing several value-sensitive scenarios and the expected impact of the described AI systems on different stakeholders. Based on the results of the sessions, we identified relevant stakeholders and values that could be impacted by the introduction of AI systems. Mapping the expected impact onto the ethical matrix gave a visual help to show clearly which values are supported by AI technology and which might be negatively affected. We have also given an example of how these values van be translated into design requirements.

The number of participants in the focus group sessions was relatively limited. Nonetheless, the participants were diverse in the function they hold in the fire organisation, which leads us to believe that the analyses give good insights into relevant values and expected impact of AI applications for fire services.

During the first focus group session we realized that to be able to identify possible impact of AI systems, a general knowledge of AI systems is needed. For that reason, we added an introduction into AI systems to the subsequent sessions. In addition, we learned that example scenarios support participants in their identification of impact. During the subsequent focus group sessions, we used scenarios to identify stakeholders and human values for fire services. The scenarios were well grounded as they were based on input from domain experts from the first focus group session, AI experts (prediction of possible technology), and literature (relevant expected values). We introduced two scenarios with different AI applications and different direct stakeholders to broaden results. The goal of the scenarios was to support the participants in thinking about AI technology and its possibilities, without limiting it to a specific implementation. The choice of these specific scenarios led to a first inventory of relevant human values, which we expect for future scenarios will be refined, extended, and if necessary amended. The chosen scenarios led mostly to the identification of impact on stakeholders within the fire services; this was intentional. In future work, we would also like to hold focus group sessions with other FR organizations (e.g., police, ambulance services) and citizens to also take their perspectives into account, as we expect that additional impact will be identified, leading to different affected values.

In the FR domain, previous research has identified relevant stakeholders and core ethically relevant themes and values for the application of rescue robots. For example, in Harbers et al.'s selection of most relevant stakeholders included victims, local authorities, electrical company, press, and observers (next to fire fighters, police, and ambulance) (Harbers et al., 2017). These stakeholders were also identified during our focus group sessions, but not prioritized for AI applications and our scenarios. For rescue robots, Harbers et al. (2017) identified the values personal safety, safety of others, access to information, well-being, effectiveness, ease of use, authority for fire fighters. Although these values do not directly correspond to the values that resulted from our sessions as we based our value definition on Friedman et al. (2013), most of them relate to them. Ease of use has not been

mentioned during our sessions. Battistuzzi et al. (2021) identified core ethically relevant themes by means of a scoping review, namely fairness and discrimination, false or excessive expectations, labor replacement, privacy, responsibility, safety, and trust. Most of these topics were mentioned during our focus group sessions; however, fairness and discrimination and labor replacement were not mentioned. During our focus groups, impact on the values *autonomy* and *identity* were mentioned clearly and were seen as very important. These values have not been identified by (Harbers et al., 2017) and (Battistuzzi et al., 2021). These values might be (more) affected by AI systems for decision support than by rescue robots, as building situation awareness and decision making are core tasks for FRs. Several human values that were identified in this research were impacted only negatively for the different stakeholders; in future research, it should be further investigated why this was the case.

In our research, some values that were mentioned by the participants correspond (partly) to some of the key requirements as promoted, for example, by the EU HLEG for AI (e.g., transparency) (AI HLEG, 2018). When comparing the values that were mentioned by the participants at the focus group sessions with the EU HLEG ethical principles and key requirements, it becomes clear that several requirements particularly relevant for AI systems were not mentioned by the participants, e.g., diversity, non-discrimination, and fairness, or environmental well-being.

There is a general consensus that AI technology raises ethical issues that are not raised by more conventional ICT technology (Floridi et al., 2018). For that reason, the original VSD list of values does not suffice for AI (Umbrello & Van de Poel, 2021) but should be supplemented by ethical principles that ensure that typical AI ethical issues are addressed. Additional to individual values, organizational and societal values and corresponding norms need to be taken into account to build socially responsible AI systems. The results of our bottom-up, empirical approach of identifying context-specific values should be supplemented by two other normative sources of values i.e., (1) values promoted by the design, such as by deriving from the UN's Sustainable Development Goals (Guterres, 2020) or from AI for Social Good principles; (2) values respected by the design, particularly values identified in relation to AI (e.g., ethical principles identified by the EU HLEG for AI (AI HLEG, 2018)).

We foresee that AI technology will not only impact the stakeholders that the participants identified as most relevant during the focus group sessions, but that AI will have impact on the first response organizations in their totality, changing roles and responsibilities, but also the way resources are allocated, how training will take place, and what is considered expertise (as is argued also for other domains, such as health care (Van Wynsberghe & Li, 2019)). Some of the impact identified by the participants supports this (e.g., expected impact on training needs, needed capabilities for first response work, and the need for a culture change). This makes it even more important to introduce AI systems responsibly, decreasing negative impact of the introduction of AI systems explicitly during the design and development process, evaluating AI systems not only in the direct interaction with the direct stakeholders, but also evaluating effects on the organizations.

The values identified in this research function as a basis for a general value set and will be iteratively built up during future research, together with a related requirements list, as well as the (related) specific values and (instantiated) requirements for critical scenarios. We started applying this research to a particular AI application for fire services (an AI-based decision-support system) (Mioch et al., 2024), to further specify values for this particular AI application, map these values onto design requirements, and operationalize them for the AI application to investigate practicability and applicability of the general framework. In addition, we will also continue with conceptual investigations by integrating values found from normative sources (such as AI for Social Good) into the results of this research, in addition to doing a scoping literature review on ethical aspects of AI systems applied to emergency response and integrating these findings into the ethical framework.

ACKNOWLEDGMENTS

We would like to thank the first responders of *Gezamenlijke Brandweer Rotterdam*, *Veiligheidsregio Rotterdam-Rijnmond*, and *Veiligheidsregio Zuid-Holland-Zuid* for their participation in this research. This research was supported by the University of Applied Sciences Utrecht (HU) through a '*promotievoucher*'.

REFERENCES

AI HLEG. (2018). Ethics Guidelines for Trustworthy AI. *European Commission*, 14–16.

Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., Hung, H., Jonker, C., Monz, C., Neerincx, M., Oliehoek, F., Prakken, H., Schlobach, S., van der Gaag, L., van Harmelen, F., . . . Welling, M. (2020). A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence. *Computer*, *53*(8), 18–28. https://doi.org/10.1109/MC.2020.2996587

Aldewereld, H., & Mioch, T. (2021). Values in design methodologies for AI. *International Conference on Advanced Information Systems Engineering*, 139–150.

Battistuzzi, L., Recchiuto, C., & Sgorbissa, A. (2021). Ethical concerns in rescue robotics: A scoping review. *Ethics and Information Technology*, *23*(4), 863–875. https://doi.org/10.1007/s10676-021-09603-0

Blasimme, A., & Vayena, E. (2020, July). The Ethics of AI in Biomedical Research, Patient Care, and Public Health. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford Handbook of Ethics of AI*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190067397.013.45

Dellermann, D., Ebel, P., So¨ llner, M., & Leimeister, J. M. (2019). Hybrid Intelligence. *Business & Information Systems Engineering*, *61*(5), 637–643. https://doi.org/10.1007/s12599-019-00595-2

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. Mit Press.

Friedman, B., Hendry, D. G., & Borning, A. (2017). A Survey of Value Sensitive Design Methods. *Foundations and Trends® in Human–Computer Interaction*, *11*(2), 63–125. https://doi.org/10.1561/1100000015

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2013). Value Sensitive Design and Information Systems. In N. Doorn, D. Schuurbiers, I. van de Poel, & M. E. Gorman (Eds.), *Early engagement and new technologies: Opening up the laboratory* (pp. 55–95). Springer Netherlands. https://doi.org/10.1007/978-94-007-7844-3_4

Galliott, J., & Scholz, J. (2020, July). The Case for Ethical AI in the Military. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford Handbook of Ethics of AI*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190067397.013.43

Guterres, A. (2020). The sustainable development goals report 2020. *United Nations publication issued by the Department of Economic and Social Affairs*, 1–64.

Harbers, M., de Greeff, J., Kruijff-Korbayova´, I., Neerincx, M., & Hindriks, K. (2017). Exploring the ethical landscape of robot-assisted Search and Rescue. *Intelligent Systems, Control and Automation: Science and Engineering*, *84*, 93–107. https://doi.org/10.1007/978-3-319-46667-5_7
cited By 9.

IEEE. (2021). Ieee standard model process for addressing ethical concerns during system design. *IEEE Std 7000-2021*, 1–82. https://doi.org/10.1109/IEEESTD.2021.9536679

IFAFRI. (2019, September). Capability Gap 9 "Deep Dive" Analysis.

Mepham, B., Kaiser, M., Thorstensen, E., Tomkins, S., & Millar, K. (2003). *Ethical Matrix Manual* (tech. rep.). LEI, Wageningen UR.

Mioch, T., Aldewereld, H., & Neerincx, M. A. (2024). Human values for responsible decision-support for fire services. In B. Penkert, B. Hellingrath, A. Widera, H. Speth, M. Middelhoff, K. Boersma, & M. Kaltho¨ner (Eds.), *Proceedings of the 21st ISCRAM conference*.

Mioch, T., Sterkenburg, R., Beuker, T., & Neerincx, M. A. (2021). Actionable Situation Awareness: Supporting Team Decisions in Hazardous Situations. *Proceedings of the 18th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2021, Blacksburg 23 May 2021 through 26 May 2021*, 62.

Morley, J., Machado, C. C. V., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *Social Science & Medicine*, *260*, 113172. https://doi.org/10.1016/j.socscimed.2020.113172

Murphy, K., Di Ruggiero, E., Upshur, R., Willison, D. J., Malhotra, N., Cai, J. C., Malhotra, N., Lui, V., & Gibson, J. (2021). Artificial intelligence for good health: A scoping review of the ethics literature. *BMC Medical Ethics*, *22*(1), 14. https://doi.org/10.1186/s12910-021-00577-8

Radianti, J., Dokas, I., Boersma, K., Noori, N. S., Belbachir, N., & Stieglitz, S. (2019). Enhancing Disaster Response for Hazardous Materials Using Emerging Technologies: The Role of AI and a Research Agenda. In J. Macintyre, L. Iliadis, I. Maglogiannis, & C. Jayne (Eds.), *Engineering Applications of Neural Networks*

(pp. 368–376, Vol. 1000). Springer International Publishing. https://doi.org/10.1007/978-3-030-20257-6_31

Schwartz, S. H. (2012). An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture*, *2*(1). https://doi.org/10.9707/2307-0919.1116

Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G. J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & So¨llner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information and Management*, *57*(2), 103174. https://doi.org/10.1016/j.im.2019.103174 Publisher: Elsevier.

Umbrello, S., & Van de Poel, I. (2021). Mapping value sensitive design onto ai for social good principles. *AI and Ethics*, *1*(3), 283–296.

van de Poel, I. (2015). Design for Values in Engineering. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains* (pp. 667–690). Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0_25

van der Stappen, E., & Steenbergen, M. v. (2020). The Ethical Matrix in Digital Innovation Projects in Higher Education. *BLED 2020 Proceedings*.

Van Wynsberghe, A., & Li, S. (2019). A paradigm shift for robot ethics: From hri to human–robot–system interaction (hrsi). *Medicolegal and Bioethics*, 11–21.

Wasilow, S., & Thorpe, J. B. (2019). Artificial Intelligence, Robotics, Ethics, and the Military: A Canadian Perspective. *AI Magazine*, *40*(1), 37–48. https://doi.org/10.1609/aimag.v40i1.2848

Wright, D., Finn, R., Gellert, R., Gutwirth, S., Schu¨tz, P., Friedewald, M., Venier, S., & Mordini, E. (2014). Ethical dilemma scenarios and emerging technologies. *Technological Forecasting and Social Change*, *87*, 325–336. https://doi.org/10.1016/j.techfore.2013.12.008

Introducing WALKER – Safeguarding Cultural Heritage through Evacuation Run Maps

Tobias Hellmund
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB
tobias.hellmund@iosb.fraunhofer.de

Jürgen Moßgraber
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB
juergen.mossgraber@iosb.fraunhofer.de

Jürgen Reuter
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB
juergen.reuter@iosb.fraunhofer.de

Susanne Rißmann
Bayerische Verwaltung der staatlichen Schlösser, Gärten und Seen
susanne.Rissmann@bsv.bayern.de

Hermann Bayer
Kreisfeuerwehrverband München e.V.
hermann.bayer.ff@gmail.com

Stefan Deschermeier
Kreisfeuerwehrverband München e.V.
stefan.deschermeier@kfv-muenchen.de

ABSTRACT

The crisis management cycle contains the step *Preparedness*, in which the development of disaster plans takes place. Run plans for fire fighters are a well-established means for enhancing preparedness in public buildings. A wide array of plans is available, serving to aid firefighters in swiftly orienting themselves during emergency situations on-site. These plans may also include detailed information as needed, providing tactical insights into critical elements within the building. In this practitioner paper, we present the adaptation of run plans to the realm of cultural heritage preservation. We utilize the existing format for the communication of information about the safe handling of cultural heritage to the command staff and the fire fighters. These *cultural heritage escape plans* offer guidance on the safe handling, evacuation procedures, and the prioritization of cultural heritage items in the case of an emergency and can be added to the existing tactical firefighting plans in public cultural heritage institutions. The evacuation plan format has been developed by the Munich County Fire Fighters and was manually implemented for all cultural heritage objects of the Bavarian Palace Department. The format has proven beneficial. With this format, we aim to enhance the preparedness of cultural heritage sites against natural disasters. Given the inherent sensitivity of cultural heritage objects, which require careful handling, these escape plans encompass instructions for the secure dismantling and evacuation of these items to designated safe areas. Furthermore, we introduce the software WALKER, a specialized tool designed for the creation of evacuation plans tailored to cultural heritage contexts. WALKER has been deployed for test usage for cultural heritage institutions for testing and evaluation of its potential. Both evacuation maps and WALKER are part of an integrated crisis management for cultural heritage and can be classified as part of the *Preparedness* step in the crisis management cycle. Furthermore, when the plans are implemented during the response, they can also be considered part of the phase *Response.*

Keywords

Evacuation Run Maps, Cultural Heritage, Software for Cultural Heritage Protection.

INTRODUCTION

Museums and historical buildings can be hit by natural disasters, just as any other building. Yet, in various workshops of the KERES project it was found, that museums and heritage preservation authorities often fail to prepare against potential natural hazards. Research in Bavaria showed 38% of cultural heritage is in natural hazardous zones and more than 400 cultural heritage objects are endangered by more than one natural hazard. All in all 136 cultural heritages in Bavaria alone were struck by natural disasters (Japelj et al. 2019). A meta study conducted by the STORM project found that a general risk strategy for cultural heritage against natural hazards in European countries is still far from reality (STORM 2018). A low-threshold measure for the establishment of a more resilient cultural heritage landscape is the creation of evacuation run maps for firefighters: evacuation maps represent an appropriate tool to enhance the safeguarding of museums and cultural heritage against threats and are a crucial component of an effective protection plan. In this paper we present the creation of run maps for the evacuation of cultural heritage objects from buildings as well as a software to manage and create this evacuation run maps. The run maps can be addended to fire brigade operation plans and form the basis of an appropriate and efficient evacuation in the case of an emergency.

TACTICAL FIREFIGHTING PLANS

The approach of emergency responders is largely standardized when there are no indications necessitating special protective measures. If firefighting needs to be conducted differently than the trained standard procedures, this information must be conveyed to the firefighters. If they lack essential information about the incident or location, their response may be less effective. Therefore, effective communication and preparation of information play a paramount role. This information must be conveyed in a standardized manner and can complement standardized firefighting procedures.

In emergency situations, it is crucial for first responders to possess a comprehensive understanding of their surroundings. The primary means of conveying vital information to firefighters is through tactical firefighting plans. These plans are customized for specific sites and serve as a communication tool between building owners and the responding firefighting teams. Typically, a firefighting plan encompasses a general overview of the building and on-site infrastructure details, including hazardous areas and water sources for firefighting. Furthermore, it incorporates tactical information, such as the requisite extinguishing agents, with the flexibility to add supplementary details as needed.

In Europe, various standards exist that delineate and specify maps for firefighters. The German standard is documented in DIN ISO 23601, which outlines the common requirements for such plans (DIN 2020). This standard is established to create a foundation in Germany and is, when necessary, elaborated upon by the regional forces. The German standard DIN 14095 delineates operation plans for firefighting and other measures on specific sites and contains information about hazards, specific characteristics and risks (DIN 2022). If additional tactical information is given, the plan is called tactical firefighting plan. The plan contains necessary information for the head of operations, such as building information, site plans or story plans or so-called *special plans*. These special plans are predestinated to hold specific information for the evacuation of cultural heritage.

Particularly in the case of water-sensitive cultural heritage items, appropriate information about the extinguishing agent are essential. The necessary containment of fire and smoke may, in some cases, not allow for the omission of water suppression. However, in these cases targeted protective measures, such as the usage of fire blankets for the covering of cultural heritage items or their evacuation, before water usage, can be initiated. The timely and targeted evacuation of cultural heritage items holds particular significance. Depending on the urgency, required personnel, and potentially necessary protective gear, evacuation measures will mostly be achievable only by firefighting personnel. However, this process cannot succeed without the support and cooperation of the facility operator: cultural heritage might be secured by protective measures, about which the firefighters need to be informed. For the purpose of planning the systematic evacuation of art and cultural assets, an initial categorization of the artworks is necessary. In this regard, the facility operator determines the order in which the cultural heritage is to be removed from the danger zone. The operator must consider the art-historical significance and the intrinsic value of the art objects. As measure to foster preparedness, the creation of tactical firefighting plans is part of emergency management for cultural heritage and can be placed in the crisis management cycle. The cycle itself discussed in the community and several definitions exist (Moßgraber 2012), yet a common definition is the subdivision into the four parts *Preparedness, Response, Recovery, Mitigation* (FEMA 2022).

Evacuation Maps for Cultural Heritage Objects

The information named in the previous chapter must be provided to the firefighters in a standardized way, so it conveys information in the right format. To address the demand for evacuation plans for cultural heritage, the Munich County Fire Department has devised a route card format, serving as an interface between firefighters and cultural heritage managers. These maps enable managers to apprise firefighters of the specifics pertaining their cultural heritage objects, thus assisting in the prevention of collateral damages. The responsibility for creating run maps is held by the managers of cultural heritage; after creating the maps these can be added to the tactical firefighting plan, as described in the previous chapter. Especially in the context of hazards due to climate change to cultural heritage, the creation of evacuation maps fosters preparedness and prevents cultural heritage loss (Kotova et al. 2023).

The Bavarian Palace Department (Bayerische Schlösserverwaltung, BSV) is a public institution entrusted with the management and preservation of cultural heritage in Bavaria, which encompasses palaces, castles, and parks. This heritage encompasses both mobile and immobile cultural items, such as historical timepieces and Neuschwanstein Castle, which houses mobile cultural heritage items. To bolster preparedness for emergencies and safeguard mobile cultural heritage, the BSV has developed evacuation maps specifically tailored to their cultural heritage objects.

WALKER – A Novel Tool for The Creation Of Run Maps

Fraunhofer IOSB developed the concepts for WALKER during the KERES project, with a primary focus on safeguarding cultural heritage against the impacts of climate change. During the project a demonstrator for WALKER has been developed. The target users for this software are cultural heritage managers trying to foster their preparedness against potential disasters endangering their cultural heritage objects. WALKER allows the creation of run maps in the format devised by the Munich County Fire Department. It simplifies the creation and updating of run maps and includes a library of predefined icons commonly used by fire departments. This ensures that the generated evacuation maps adhere to recognized standards, making them practical for real-life applications. Furthermore, WALKER offers the evaluation of stored data, to allow the quick retrieval cultural heritage objects fitting a specific criterion, such as priority or the number of personal required for the transport.

WALKER has been developed in the context of the KERES project, which was accompanied by an expert panel. Members of the panel were cultural heritage institutions, as well as authorities and organizations with civil security tasks, such as fire fighters or the federal technical aid organizations. WALKER was developed together with the expert panel, who reviewed the application and requested or suggested new features. As the development took place during the COVID19 pandemic, the feedback sessions were held online in which the application was demonstrated. For interested users a testing environment was setup, which they could access for testing purposes. Suggestions and functions were defined in use cases, from which the most promising were chosen to be integrated into the prototype application. The development process followed the waterfall project management paradigm in four phases. Together with the experts a specification document for the software was developed. The specifications were solved with a draft, that offered (to that point hypothetical) functions, which were mapped to the requirements from the first phase. The third phase comprised the implementation of the developed functions. Quality assurance formed the last phase, in which software and end users tests were conducted.

The application allows the creation of evacuation maps through forms. Figure 1 shows the view used to create and update the evacuation maps. Information, such as position, priority or inventory numbers can be entered; images can be uploaded additionally.

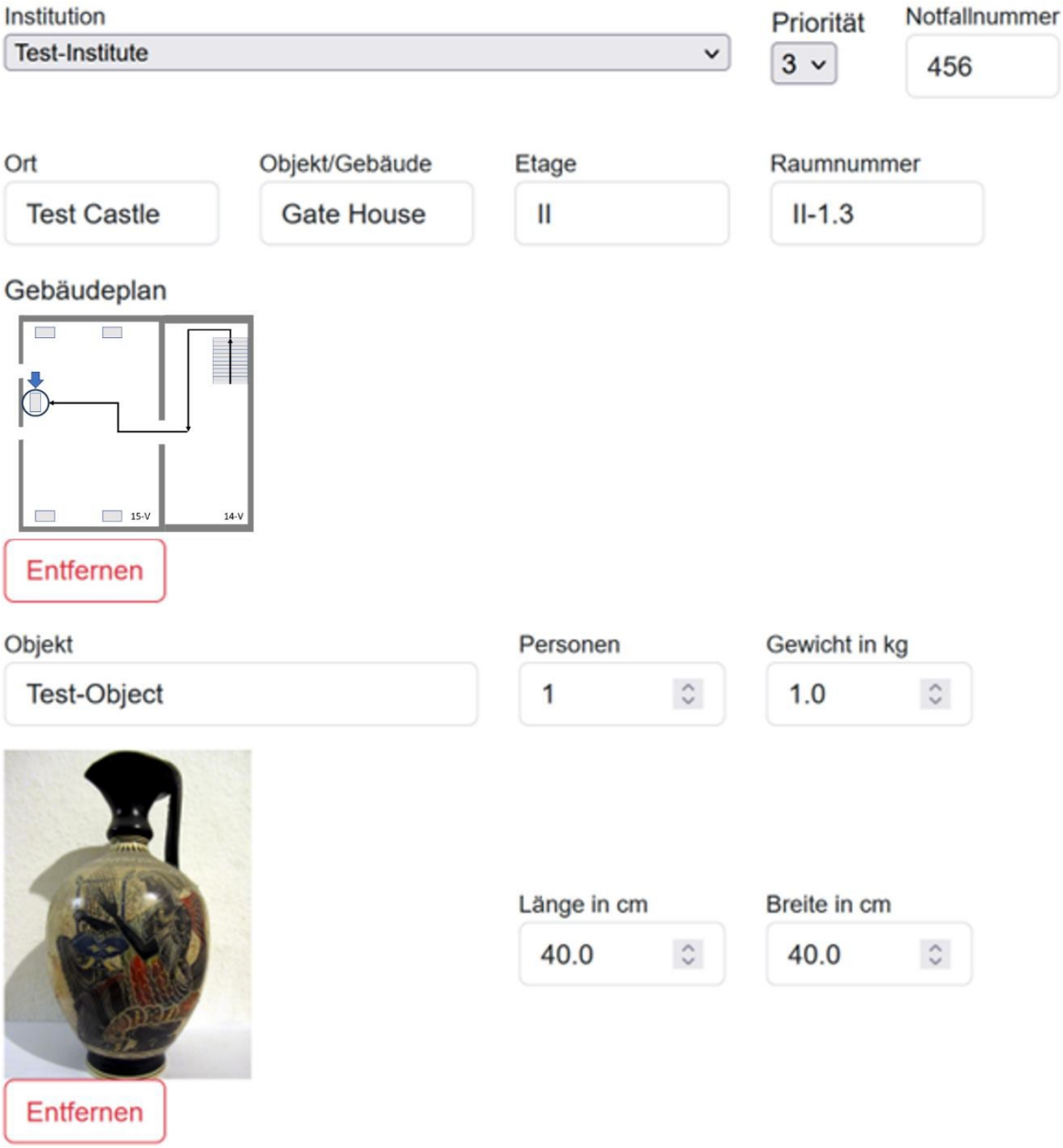


Figure 1. An excerpt from the WALKER form for data input

After entering the data, users can trigger the generation of the corresponding PDF-files which are offered as download. A created evacuation map is shown in Figure 2. It shows a hypothetical cultural heritage object managed by an institute called Test-Institute. The map was created on 09/08/2023 by a user called Walker. It has the priority level 3. The position of the object is shown in the map with an arrow; an image is shown as well. The object weighs one kilogram and measures 40 x 40 x 70 cm. A single person is required to transport the object. An image (broken vase) signalizes the object needs to be handled with care. This route card could be added to the tactical firefighting plan, thus supplementing the existing information for the firefighters.

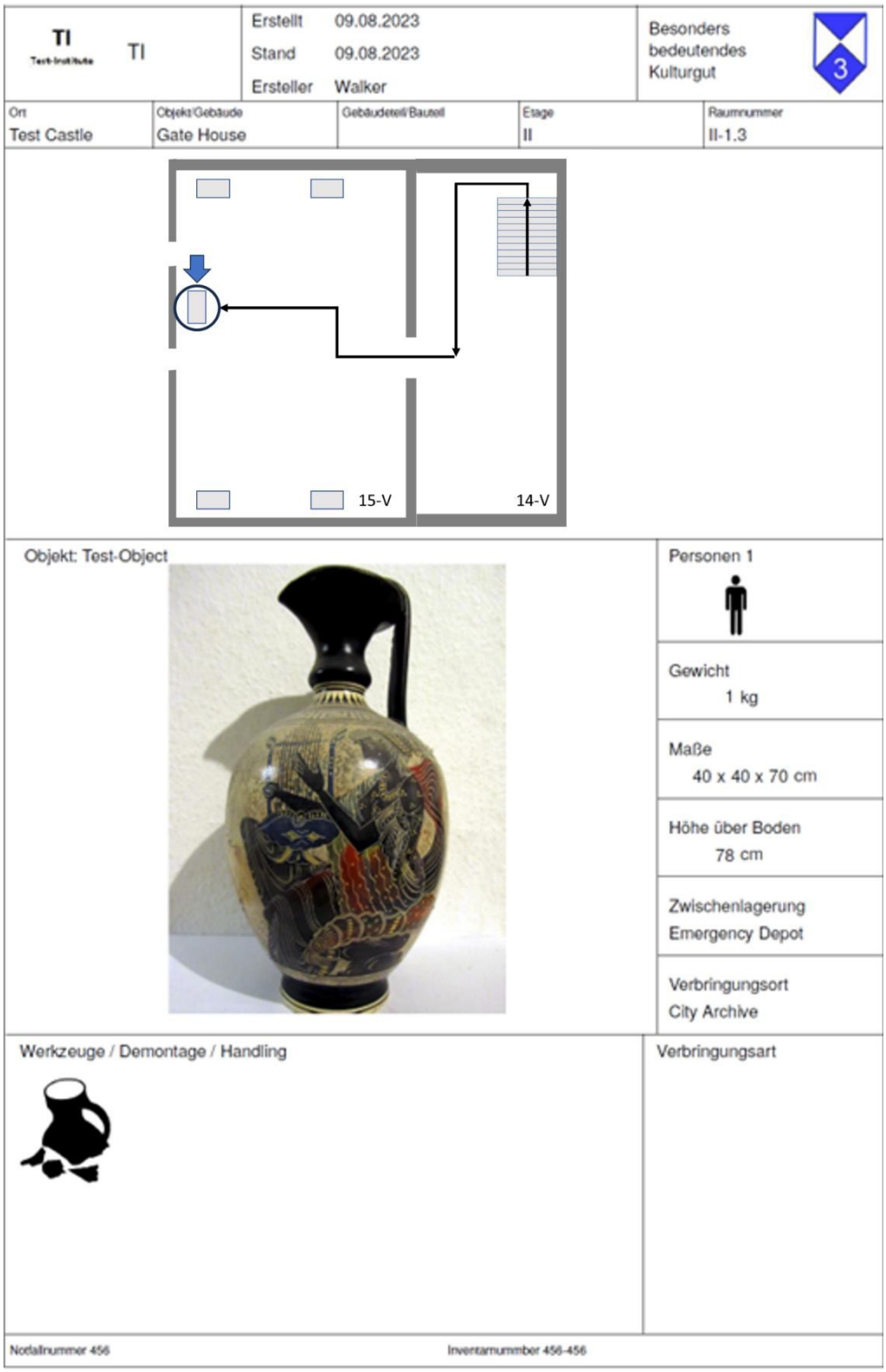


Figure 2. The WALKER generated route card

In the workshops, further beneficial features were pointed, such as the ability to filter evacuation maps by priority level and the required resources for its evacuation to be able to prioritize the evacuation of objects in case of an emergency. Yet, features like these face the contrary demand of availability of these maps in a printed form by authorities and organizations with civil security tasks. This aspect could be overcome by utilizing WALKER in the crisis operations center as application: the head of operations could, for example, request route maps with the highest priority or query objects that need only the firefighter personal he has available.

WALKER – ARCHITECTURE

WALKER is a web application built on Java 17. It is built on a backend developed on the Spring Boot framework and a front-end implemented as a Vue 3 single-page application, with Bootstrap used for layout purposes. Communication between the front-end and backend takes place through an HTTP programming interface. Structured data in WALKER is stored in a PostgreSQL database, although it can potentially work with any SQL-based database. Unstructured data, such as images, is stored within the file system. The file hierarchy is designed in a way that files from different institutions are stored in different locations if the file system. The database maintains records of file locations to ensure the unique identification of all necessary objects. Creating or updating data in WALKER can be accomplished in two ways: Users have the option to input all the necessary data for creating a route card through web forms. These forms also facilitate data updates, so the user uses the same forms for creating and updating data. The backend offers a REST-API, allowing for the (semi-)automatic creation of maps by uploading datasets. REST outlines principles for services without specifying implementation details. The first principle is the Client-Server Architecture, as seen in the World-Wide-Web, where a server provides a service that clients can request. Utilizing the widely adopted HTTP protocol (Internet Engineering Task Force 1999), REST client implementations are accessible across various programming languages. The second key characteristic is statelessness, requiring every message sent to a REST service to include all necessary information for processing. The common HTTP-methods are used to interact with the interface, whereas GET is used to retrieve entities, POST creates new entities, PUT and PATCH update entities and DELETE finally deletes entities (Roy Thomas Fielding 2000). WALKER employs LaTeX for map generation, which is triggered through the Java backend. The maps are made available as Portable Document Format files (PDF). Parameters, including required images and textual information, are stored in a configuration file. This file is then passed to as command line parameter to LaTeX, which, in turn, generates the map in PDF format. Passing the file as parameter should prevent command line injections. Assessment for potential publication on well-known open-source repository management platforms, such as GitHub, is ongoing.

EVALUATION

For further usability testing, WALKER has been deployed for BSV, as well as for Leibniz-Institut für Biodiversitätsforschung and will be deployed for Staatliche Schlösser und Gärten Baden-Württemberg. For deeper insights, we aim to evaluate both the practicability of WALKER and the user experience. An important aspect is the alignment of requirements due to different standardized operations in the federal states in Germany: differences in law and operations might have an influence on the usage of WALKER and its integration into the common workflows of a cultural heritage institution. WALKER has generated curiosity in the cultural heritage community, as it offers a low-threshold solution for the creation of evacuation maps and to enhance the preparedness of cultural heritage institutions. All test-users assess the software regarding its applicability, most importantly the question is, whether WALKER can provide benefit in a real-world context. For this, WALKER needs to undergo usability tests and most importantly expert reviews, with experts from the firefighting domain, as well as cultural heritage managers to find potential shortcomings. The created route cards could also be assessed through simulations and expert reviews, but as they are designed by fire fighters, their usability should be promising. In a dedicated upcoming project, we envisage to continue test and develop WALKER. The run maps have been evaluated during a test with fire fighters at Castle Schleißheim: before the drill, the handling and evacuation for all 400 objects within the castle have been documented with run maps in the described format (Bachhuber 2017). Feedback had been implemented in the route card format, which again was incorporated into the WALKER software.

CONCLUSION

In this paper we presented the transferal of run maps to the domain of cultural heritage for the evacuation of cultural heritage objects in the case of an emergency. By utilizing a well-known format of maps, we lower the acceptance threshold and foster practicability and the ability to integrate these maps into the common operational procedures. The format was developed by the Munich County Fire Fighters and was made publicly available for fire fighters and the managers of cultural heritage. The format has been adopted by the Bavarian Palace Department, who manually created evacuation maps for all relevant objects in their collections and assets. The maps have been tested in an evacuation drill with 400 objects in Castle Schleißheim. In the run of the KERES project, Fraunhofer IOSB developed a prototype software for the software-supported creation of maps in the named format. The software was deployed for several expert users. Further development and testing is envisaged in an upcoming project. The tests should evaluate the software and its application itself, but also its integration into the security framework of the using institutions.

REFERENCES

Bachhuber, K. (2017). Höchste Auszeichnung. In: Süddeutsche Zeitung, 01.10.2017. Available online: https://www.sueddeutsche.de/muenchen/landkreismuenchen/unterschleissheim-hoechste-auszeichnung-1.3690725, last accessed 11/16.2023.

DIN (2022). DIN 14095. Feuerwehrpläne für bauliche Anlagen. Available online https://www.din.de/de/mitwirken/normenausschuesse/fnfw/entwuerfe/wdc-beuth:din21:358653586, last updated 11/16/2023, last accessed 11/16/2023.

DIN (2020). DIN ISO 23601. Sicherheitskennzeichnung - Flucht- und Rettungspläne. Available online: https://www.din.de/de/mitwirken/normenausschuesse/nasg/veroeffentlichungen/wdc-beuth:din21:343877361, last accessed 11/09/2023.

Federal Emergency Management Institute (2002). Emergency Management in the United States. FEMA. Emmitsburg, Maryland, USA. Available online https://training.fema.gov/emiweb/downloads/is111_unit%204.pdf, last accessed 11/16/2023.

Internet Engineering Task Force (1999). RFC 2616: Hypertext Transfer Protocol -- HTTP/1.1. Available online https://datatracker.ietf.org/doc/html/rfc2616, zuletzt aktualisiert am 16.11.2023, last accessed 11/16/2023.

Japelj, A., & Ferreira, A., & Planinšek, S., & Pintar, A. (2019). Alpine cultural heritage and natural hazards: where do we stand? : review and survey - framework analysis on Alpine cultural identity and natural hazards. Available online: https://dirros.openscience.si/izpisgradiva.php?id=10979, last accessed 3rd November 3, 2023.

Oracle Corporation (2023). Java Platform, Standard Edition 17 Reference Implementations. Available online https://jdk.java.net/java-se-ri/17, last updated 22.09.2023, last accessed 16.11.2023.

Kotova, Lola; Leissner, Johanna; Winkler, Matthias; Kilian, Ralf; Bichlmair, Stefan; Antretter, Florian et al. (2023). Making use of climate information for sustainable preservation of cultural heritage: applications to the KERES project. In: *Herit Sci* 11 (1), S. 1–18. DOI: 10.1186/s40494-022-00853-9.

Moßgraber, J. (2017). Ein Rahmenwerk für die Architektur von Frühwarnsystemen (Vol. 29). KIT Scientific Publishing.

Roy Thomas Fielding (2000). Architectural Styles and the Design of Network-based Software Architectures. University of California. Irvine, CA, USA. Online verfügbar unter https://ics.uci.edu/~fielding/pubs/dissertation/top.htm, zuletzt aktualisiert am 15.03.2002, zuletzt geprüft am 09.11.2023.

STORM (2018). D2.1 - State of the art policies on government of Cultural Heritage against natural disasters and climate changes. Available online: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5bb1c5ec0&appId=PPGMS, last accessed 11/09/2023.

Vielhuber, J (2015). INFO Kulturgutschutz. Für Einsatzkräfte der Feuerwehren im Landkreis München und für Betreiber von Kunst- und Kulturstätten. Published by Kreisfeuerwehrverband München e.V. Available online: https://www.lfv-bayern.de/media/filer_public/13/22/132274c5-e49c-4b77-bf8b-4a01bec4eef2/29-information_kulturgutschutzplan_2015_12_23.pdf, last accessed 11/16/2023.

The Utility of Community Resilience Models for the Prevention of Online Violence

Dante Della Vella
Vestlandsforsking
ddv@vestforsk.no

Hoang Long Nguyen
Vestlandsforsking
longnh238@gmail.com

Rajendra Akerkar
Vestlandsforsking
rak@vestforsk.no

ABSTRACT

Online violence is a complex issue transcending national borders with digital and real-world consequences. It is a consequence of hate, fueled by real and perceived injustices. Communities must be able to adapt - improving citizens' social and material conditions to prevent hate from taking root or spreading. Community resilience research focuses on assessing a community's ability to adapt and how well it copes with disruptions both known and unknown. Therefore, community resilience models and methods should be adopted by those working to prevent online violence both to investigate new means of preventing hate speech and to test existing community resilience theory in a new context. This paper will review online violence and community resilience, explain their connection, review community resilience models that may be of interest for violence prevention work, and describe how those models are being used in a project that is currently in progress.

Keywords

Community Resilience, Hate Speech, Online Violence, Violence Prevention, Sustained Adaptability.

INTRODUCTION

Communities of people are complex adaptive systems comprising individuals, social networks, technologies, and the environment. To thrive over time, all complex systems must adapt, changing their underlying architecture to meet their goals as the external environment exposes them to new disruptions. For a community, this means maintaining citizens' quality of life, accommodating new peoples and technologies, and preserving the natural environment. The capability to continually adjust over time has been labeled Resilience (Hollnagel et al., 2006). Resilience is an emergent property fostered by the interaction of components and subsystems such that the whole system is poised to adapt and continues to function as conditions change over time (Woods, 2015). Communities therefore constantly engage in work to make sense of the world, plan for the future, and coordinate resources to maintain a high quality of life for citizens. Taking steps to foster resilience will allow communities to perform this work while dealing with the shocks and threats of the world.

One of the most urgent threats of the present day is the proliferation of hate speech online. Hate speech is an example of online violence - a continuum of behavior from subtle discrimination to violent crime (Gagliardone et al., 2015; Kim, 2021.; Lupu et al., 2023). Though hate speech is not new, the internet - because of its transnational reach, accessibility, and the anonymity it permits provides an excellent medium through which hate speech may spread (Gagliardone et al., 2015; Kim, 2021). Research has shown that hate speech creates real tangible threats of violence to citizens, particularly the most vulnerable members of a community. Though online and offline phenomena are substantively related their relationship is complex and still not well understood (Lupu et al., 2023). While the need to address hate speech online is clear, the practice of addressing it proves complicated. It requires

communities to navigate fundamental tradeoffs between protecting individual liberties and the safety of citizens (Gagliardone et al., 2015). Limitations on speech must be specified in law, proportionate, and targeted to be justifiable. Victims are heterogeneous – people may become targets based on race, ethnicity, sexual orientation, gender identity or expression, religion, or disability. Furthermore, these behaviors have structural, interpersonal, and individual pre-determinants and impacts; effective prevention of violence-inducing behaviors requires an in-depth understanding of those pre-determinants. In other words, an understanding of the material and social characteristics of a community (its underlying architectural properties), how they interact to enable the spread of hate speech, and how they are affected by its spread and consequences. These principles are precisely the same as those central to research on community resilience.

The purpose of this paper is to elucidate the relevance of community resilience research to work preventing online violence such as hate speech. In section II we will elaborate on the concepts of online violence, hate speech, and community resilience and the links between them. In section III several models from the community resilience literature will be compared. In section IV we will discuss how these models could be applied to online violence prevention work and will recommend the most promising candidates for this application.

This work is ongoing and is part of the SOCYTI project - an in-progress, transdisciplinary initiative dedicated to developing new strategies and tools for the prevention of hate speech online. The work described in this paper represents the initial stages of our efforts to carry out this integration of community resilience research into online violence prevention. The project aims to develop a cloud-based, real-time detection system capable of evaluating text and images from social media posts for hateful content on a larger scale than ever previously possible. It is the hope of the research team that such tools can enable proactive interventions rather than just being used to reactively police hate speech online. Community Resilience research is integral to realizing this hope. In the future we will directly connect community resilience indicators to specific types of hate speech. These connections may reveal new means of preventing different kinds of hate speech from propagating.

LITERATURE REVIEW

Online Violence and Hate Speech

Hate speech is a product of hatred - a feeling of intense hostility and aversion, usually deriving from fear, anger, or a sense of injury. It relies on underlying tensions in society, which it seeks to re-produce and amplify. It is the expression of conflicts between different groups (Gagliardone et al., 2015; The Rabat Plan of Action, 2012). It is intended divide citizens into groups of 'us' and 'them' and then blame 'them' for the struggles of 'us'. Such scapegoating obfuscates the real sources of threats to communities. It degrades the community's capacity to monitor, anticipate, and respond to real threats, which are key cornerstones of system resilience (Hollnagel, 2009). This systemic resilience extends beyond just hate speech, dealing with how complex systems cope with all kinds of disruptions including natural and man-made disasters. As such, hate speech represents a legitimate compounding threat for myriad crises and disasters. In addition, such events in and of themselves can contribute to and exacerbate the aforementioned social and economic tensions which feed into the propagation of xenophobic ideology and sentiment.

Hatred can be constructed along several axes including race, ethnicity, language, gender identity or expression, religion, sexuality, or nationality. It may stem from political polarization – 'having an enemy to oppose provides people with a sense of identity, and the possibility of victory in the face of issues that seem too large to hold' (The Rabat Plan of Action, 2012). Economic issues such as unemployment, declining income, or increased living costs (especially during recessions) often contribute to anti-migrant sentiments. Groups that are excluded and/or discriminated against can become radicalized, which diminishes the possibility of establishing cohesion and cultural dialogue (Eminoğlu, 2020). Any combination of these factors, along with the absence of free and secure civic space, online disinhibition, and more may drive radicalization and the hate speech that results from it (The Rabat Plan of Action, 2012; UNOCT PVE Reference Guide, 2016) hence we label them *drivers*. The above list is not exhaustive. The structure and context of a community determine which drivers are present and how they combine to foster online violence.

Not all hate speech is illegal. 'Only the most egregious forms of hate speech, namely those constituting incitement to discrimination, hostility, and violence, are generally considered unlawful (Gagliardone et al., 2015). While legal responses such as protections for migrant workers and diaspora communities are important, legislation is only one of many tools. legislation must be complemented by initiatives from various sectors of society geared towards the deeper structural issues that foster radicalization, hatred, and hate speech – the drivers. Addressing such issues will require a plurality of policies, practices, and measures nurturing social consciousness, tolerance, and understanding (Kim, 2021; UNOCT PVE Reference Guide, 2016). Strong community building initiatives,

education and training programs, and intercultural dialogue foster acceptance and cultural literacy through interaction. Improving socio-economic parity, ensuring equitable access to basic and specialized services, and dedicating resources towards anti-radicalization plans assuage tensions on which hate speech depends (Eminoğlu, 2020). Healthy cooperation between local actors, civil organizations, and government is important to ensure that such changes are well supported and spread throughout society.

Community Resilience and Models

The word 'community' can describe groups united by common interest, geography, or socioeconomic, cultural, political, or ethnic makeup. Community resilience describes the capability of a community, enabled by an interconnected network of systems, to continuously meet its primary goals to keep citizens healthy, happy, and safe (Fitzpatrick, 2016). These systems comprise socioeconomic, ecological, and built aspects. A community is resilient when: '… the population are connected to one another and work together, … to function and sustain critical systems, even under stress; adapt to changes in the physical, social, or economic environment; be self-reliant if external resources are limited or cut off; and learn from experience to improve over time' (Fitzpatrick, 2016). We have identified research investigating these underlying concepts from many sources using the labels 'community resilience', 'societal resilience', and 'social resilience' and will treat each label as synonymous in this review.

Synthesizing existing strands of community resilience literature, the first from social-ecological systems and the second from psychology, yields the following set of core principles: People–place connections; values and beliefs; knowledge, skills, and lea rning; social networks; engaged governance; diverse and innovative economy; Community infrastructure; leadership; and readiness to accept change (Berkes & Ross, 2013). The agency of actors in the community and the process of self-organization both emerge from the state and combination of these principles. *Fostering agency in people* and *allowing for self-organization* are thus important foci for efforts to build community resilience.

Societal resilience literature suggests that 'societal resilience lies in techno-political macrostructures that actively shape societies and risks at a global level as well as in the skills of professionals' (Haavik, 2020). Techno-political macrostructures refer to critical infrastructures (power, telecoms, water, etc.) and critical societal institutions (health, political, family, etc.). This perspective contrasts with traditional safety science, which focuses wholly on infra structure robustness and restricts the ability to explore the nature of risks that transcend infrastructures and national borders such as the proliferation of hate speech online. 'If we accept critical infrastructures as the central risk object, we indirectly consolidate the politics, the technologies, and the practices that admit these structures their criticality' (Haavik, 2020). Adopting the new perspective creates space to investigate the adaptive capacity of a society by centering the socio-ecological production of risks.

Some social resilience literature argues that social resilience builds on the general concept of resilience, rooted in ecological systems thinking, and departs from it by incorporating an actor-oriented perspective recognizing the importance of power, politics, and participation. This conception of social resilience comprises three dimensions: Coping Capacities - the ability of social actors to cope with and overcome adversities, Adaptive Capacities - their ability to learn from past experiences and adjust to future challenges, and Transformative Capacities - their ability to craft sets of institutions that foster individual welfare and sustainable societal robustness towards future crises (Keck & Sakdapolrak, 2013).

A community resilience model must incorporate sociological, ecological, and infrastructural elements, an inherently complex task leading to many different model forms. Models can be qualitative, quantitative, or hybrid. Qualitative models create theoretica l frameworks, highlighting key components of resilience and their schematic connections. Quantitative models generally seek to make resilience comparable between entities and locations, and trackable across time (Scherzer et al., 2019). A common modeling strategy is to deconstruct a community into 'domains.' Domains represent functional areas inherent to a standard community that can be assessed for resilience. Sometimes these are called 'dimensions,' or 'subdomains.' Qualitative or qua ntitative indicators are then organized into domains. Several domains reappear across models and others are functionally identical while being labeled slightly differently. Some models include unique domains or are not organized into domains at all. In the following section, we will inspect several models and frameworks from the literature. This paper will cover five models, two frameworks, one toolkit, and one strategy for combatting cyber-racism. We will group them according to how their indicators are structured. A summary of these sources is shown in Table 1.

Table 1. Overview of Sources from Community Resilience Literature

Name	Type
Resiloc Resilience Indicators	Model
Disaster Resilience Indicators for Benchmarking Baseline Conditions (BRIC)	Model
Community Resilience Index for Norway (CRIN)	Model
City Resilience Index (CRI)	Model
Communities Advancing Resilience Toolkit (CART)	Toolkit
Community Disaster Resilience	Framework
Cyber Racism and Community Resilience	Strategy
The Community Capitals Framework	Framework
Intercultural Cities: Intercultural Dialogue as a Tool for Social Cohesion	Model

COMPARATIVE ANALYSIS

The RESILOC Model, The Disaster Resilience Indicators for Benchmarking Baseline Conditions (BRIC) Model, and the Community Resilience Index for Norway (CRIN) model are highly similar structurally. We label this shared structure Domain Set 1 and visualize it in Figure 1. All three contain *Social*, *Infrastructure*, and *Economy* domains. The *Social* domain captures general demographic characteristics of a population that enhance their ability to deal with adverse events. It includes relationships between people and groups, social networks, and culture (Scherzer et al., 2019; Cutter et al., 2010; Akerkar et al., 2022). The *Infrastructure* domain relates to physical structures, technical facilities, and systems essential to community functioning. These include transportation, electricity, gas, water, communications systems, food distribution chains, health facilities, municipal services, and public administration services. The *Economy* domain intends to portray overall economic health and thus focuses on indicators such as variability and diversity of income sources, range and types of local businesses, assessments of existing wealth, and the income generating skills-base of citizens (Scherzer et al., 2019; Cutter et al., 2010; Akerkar et al., 2022).

RESILOC uniquely contains a domain for *Government* and *Disaster Risk Reduction*. BRIC and CRIN both contain a domain for *Institutions* and *Community Capital*. RESILOC and CRIN both contain an *Environment* doma in. The *Government* domain and the *Institutional* domain are both concerned with institutional efficacy and transparency, and citizen engagement in political processes. Because the indicators present in each are so similar, we treat the two domains as synonymous. Thus, we visualize them as a single domain as shown in figure 1.

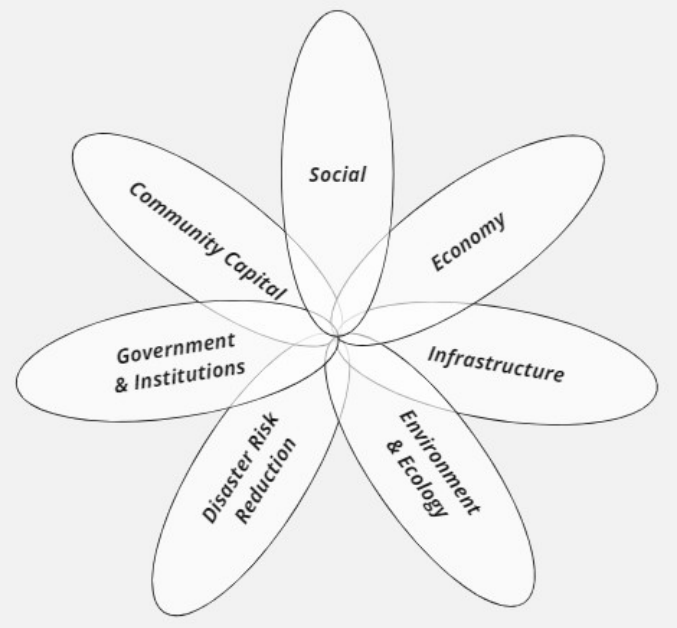


Figure 1: Domain Set 1

The indicators of the *Government/Institutions* domain often overlap with other domains because institutions play an important role in organizing and regulating social, political, and economic activities (Akerkar et al., 2022). Indicators within the *Infrastructure* domain also often overlap with other domains because physical infrastructures can serve purposes specific to the other domains (Akerkar et al., 2022). We keep the *Infrastructure* domain distinct to maintain space for assessing the purely physical attributes of infrastructure such as infrastructure quality, or the proximity of citizens to important infrastructure (Scherzer et al., 2019; Cutter et al., 2010).

The *Community Capital* domain is highly related to the *Social* domain but is unique in seeking to capture capacities such as the strength of community organizations and citizen knowledge, skills, and resources rather than social relationships and connection between people (Scherzer et al., 2019; Cutter et al., 2010). Across reviewed models there is considerable indicator overlap and potential misplacement between the *Social* and *Community Capital* domains. Despite this, a conceptual distinction between indicators that point towards social relationships and community capital is important, so we keep these domains distinct.

The *Disaster Risk and Reduction* domain describes frameworks, processes, and resources dedicated to risk reduction and emergency management. This includes equipment, temporary housing, statutory and voluntary organizations, etc. (Akerkar et al., 2022). Though many such indicators are intrinsically linked to the government of a community this domain is kept distinct to accentuate a focus on processes and capacities specifically for crisis situations.

The *Environment* domain has two focuses. First, assessment of the natural environment that intersects and envelopes a community (i.e., hazard and exposure mapping of local topography, geography, geology, vegetation, and biodiversity). Second, assessment of a community's exposure and resilience to natural hazards (Akerkar et al., 2022). Such indicators are sometimes organized as the *Ecology* domain. Because the indicators of both these domains are environmentally focused, we treat them as synonymous. Conversely, while there is also considerable overlap between *Environment/Ecology* domain and *Disaster Risk Reduction* domain indicators, we keep these domains distinct to recognize that communities must prepare for man-made as well as natural disasters and to maintain space for environmental indicators not directly related to crises.

Some models group indicators into an *Individuals*, or *Physical*, or other similarly narrow domain (Nguyen & Akerkar, 2020). This review is interested in models which consider resilience as a systems-level property, not as something that resides in individual citizens alone. To prevent losing the value of these indicators we recommend that these domains be disregarded, and their indicators reorganized into another domain present in that model as appropriate. For example, the community domain (human capital/knowledge, skills & resources) or the infrastructure domain (physical structures).

The City Resilience Index (CRI) and the Communities Advancing Resilience Toolkit (CART) represent other forms of organization. The CRI comprises four dimensions split into twelve goals into which indicators are grouped (City Resilience Index, 2015). We label this structure Domain Set 2 and present a visualization in Figure 2.

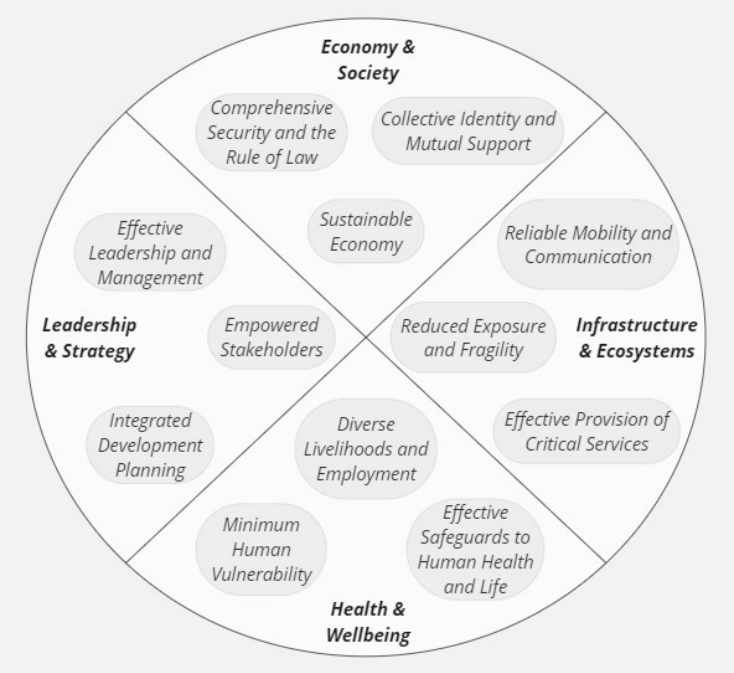


Figure 2: Domain Set 2

The *Leadership and Strategy* dimension includes indicators that would be found in the *Government/Institutions*, *Community Capitals*, and *Disaster Risk Reduction* domains of domain set 1. The *Infrastructure and Ecosystems* dimension includes indicators that would be found in the *Infrastructure*, *Environment/Ecology*, *Government/ Institutions*, and *Disaster Risk Reduction* domains. The *Economy and Society* dimension contains indicators that would be found in the *Economy*, *Social*, *Community Capitals*, and *Government/Institutions* domains. The *Health and Wellbeing* dimension contains indicators that would be found in the *Disaster Risk Reduction*, *Economy*, and *Infrastructure* domains. A summary of the comparison of domain set 1 and domain set 2 is shown in Table 2. Domain set 1 is shown in the first column. The presence of an X in a table cell indicates that the domains listed in that column and row contain analogous indicators as described above.

Table 2. Comparison Between Domain Sets 1 and 2

Domain (Set 1)	**Leadership & Strategy**	**Economy & Society**	**Infrastructure & Ecosystems**	**Health & Wellbeing**
Government & Institutions	X	X	X	
Community Capitals	X	X		
Social		X		
Economy		X		X
Infrastructure			X	X
Environment & Ecology			X	
Disaster Risk Reduction	X		X	X

CART does not contain specific indicators and only describes four overlapping domains of resilience. These are *Connection and Caring*, *Resources*, *Transformative Potential*, and *Disaster Management* (Pfefferbaum et al., 2013). We label this organizational structure Domain Set 3 and present a visualization in figure 3.

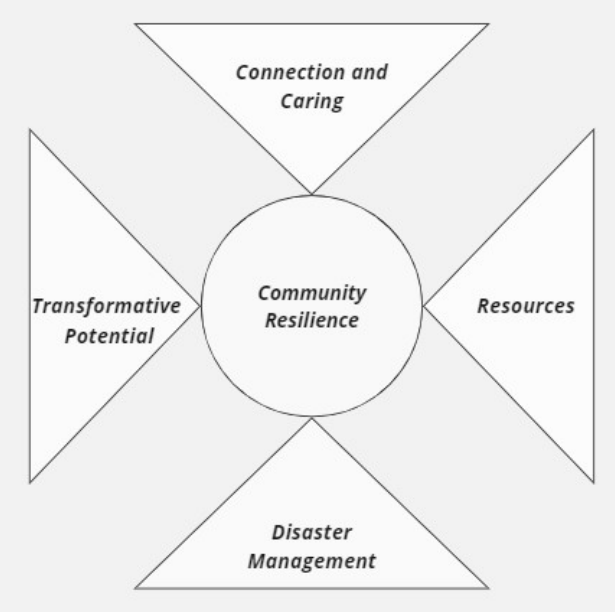


Figure 3: Domain Set 3

Connection and Caring focuses on social relationships, support systems, and the connection between people, which would be found in the *Social* domain. Resources includes natural, physical, information, human, socia l, and financial resources, which encompasses the *Environment*, *Community Capital*, *Economy*, and *Infrastructure* domains. *Disaster Management* concerns everything related to prevention, mitigation, preparedness, response, and recovery for crises. These indicators are closest to the *Disaster Risk Reduction*, *Environment/Ecology*, and *Government/Institutions* domains of the previous models. Finally, *Transformative Potential* is the model's most unique domain. It describes the ability to identify and frame collective experiences, examine successes and failures, and engage in critical analysis. Such indicators would be found in the *Government/Institutions*, *Community Capitals*, *Social*, and *Environment/Ecology* domains. A summary of the comparison of Domain Set 1 and Domain Set 3 is shown in Table 3 and a comparison between Domain Sets 2 and 3 is shown in Table 4.

Table 3. Comparison Between Domain Sets 1 and 3

Domain (Set 1)	**Transformative Potential**	**Connection & Caring**	**Resources**	**Disaster Management**
Government & Institutions	X			X
Community Capitals	X		X	
Social	X	X		
Economy			X	
Infrastructure			X	
Environment & Ecology	X		X	X
Disaster Risk Reduction				X

Table 4. Comparison Between Domain Sets 2 and 3

Domain (Set 2)	**Transformative Potential**	**Connection & Caring**	**Resources**	**Disaster Management**
Leadership & Strategy	X	X		X
Economy & Society		X	X	
Infrastructure & Ecosystems	X		X	X
Health & Wellbeing			X	X

Several sources do not organize their insights into a structured model. The *Community Disaster Resilience* (Fitzpatrick, 2016) framework provides a broad set of principles that explain why some communities cope with disaster events better than others. The paper on *Cyber Racism and Community Resilience* (Jakubowicz et al., 2017) positions racism and other forms of hate speech online as one of many challenges that undermine social cohesion and break down social capital. It also identifies resilience as the most critical element in resisting and overcoming hate speech. The *Community Capitals Framework* (Pigg et al., 2013) describes seven types of capital present in communities on which development efforts can focus. *Intercultural Cities: Intercultural Dialogue as a Tool for Social Cohesion* (Eminoğlu, 2020) explores a new model for social cohesion – interculturalism – and reports key aspects. These sources also show considerable overlap in their indicators.

All these sources contain indicators that relate to government, social networks and relationships, and community capital. Many of these indicators are written differently but point towards the same phenomena such as existing policies, civil engagement in political processes and community organizations, social cohesion and mutual care, and the skills of individuals. Indicators focused on the economy and the natural environment are present in the *Community Capitals Framework* and the paper on *Community Disaster Resilience*. Indicators focused on infrastructure are present in the *Community Capitals Framework* as well as the papers on *Cyber Racism and Community Resilience* and *Intercultural Cities*. Indicators focused on disaster risk reduction are present in the papers on *Community Disaster Resilience* and *Cyber Racism and Community Resilience*. A summary of the comparison of Domain Set 1 and these individual sources is shown in Table 4. A comparison between Doma in set 2 and these sources can be inferred by referencing table 2. A comparison between Domain Set 3 and the unorganized sources is shown in table 5. Both *Community Disaster Resilience* and *Cyber Racism and Community Resilience* provide indicators that fit into each domain of set 3. *Intercultural Cities* presents indicators that are related to connection and caring and resources. Interestingly, *The Community Capitals Framework* only provides indicators that are categorizable as resources.

Table 5. Comparison Between Domain Set 1 and Unorganized Sources

Domain (Set 1)	**Community Capitals Framework**	**Community Disaster Resilience**	**Cyber Racism & Community Resilience**	**Intercultural Cities**
Government & Institutions	X	X	X	X
Community Capitals	X	X	X	X
Social	X	X	X	X
Economy	X	X		
Infrastructure	X		X	X
Environment & Ecology	X	X		
Disaster Risk Reduction		X	X	

Table 6. Comparison Between Domain Set 3 and Unorganized Sources

Domain (Set 1)	**Community Capitals Framework**	**Community Disaster Resilience**	**Cyber Racism & Community Resilience**	**Intercultural Cities**
Transformative Potential		X	X	
Connection & Caring		X	X	X
Resources	X	X	X	X
Disaster Management		X	X	

DISCUSSION

Regardless of form and organization, all the models shown exhibit commonalities. The health and well-being of citizens is paramount. Attention is repeatedly paid to the extent to which all people can meet their basic needs, particularly in times of crisis. The community must be able to provide for the health and well-being of citizens and safeguard it through normal and emergency situations. This capacity may be captured via healthcare facilities, public health initiatives, economic structure, governmental policy, environmental health, etc. Attention is also paid to the organization of communities. Indicators cover culture and social structure as well as the makeup of the built and natural environment. Assessment of the quality of that environment is also seen across models. A comprehensive resilience perspective requires each element because social structures inform the kinds of infrastructure that is prioritized, and the existing built and natural environment shapes social organization, in addition, different kinds of infrastructure and ecosystems provide protection from some disasters and create vulnerability to others. The models are also concerned with the continuity of natural and infrastructural resources when faced with shock. Finally, the models capture knowledge and skills across layers in the community network. Resilient communities learn from the past and take steps to anticipate the future by making sense of the world around them. The knowledge that enables this is held by governmental bodies and institutions, as well as by citizens themselves, perhaps clustered into civil society organizations. What knowledge and skills exists, who holds that knowledge and those skills, and how to ensure those holders are included in decision making, is all critical for building the resilience of a community.

The relationship between violence inducing behaviors such as online hate speech and community resilience is a positive feedback loop. People are taught to hate. Communities with low resilience struggle to maintain their critical functions of keeping citizens happy, healthy, and safe. This creates tensions that can drive the spread of hatred. The exact combination of drivers that are present in one community changes according to what functions are failing or vulnerable. Hatred can spread through online platf orms rapidly, causing some groups to be ostracized and discriminated against. Discriminated people are dissuaded from participating in civic life which robs the whole of their perspective and skills and reduces their access to important resources needed to maintain their health and preparedness for emergencies. The separation of different groups degrades social cohesion making it more difficult for knowledge to transfer throughout the system, and inequitably exposes certain groups to dangers while reducing their means to cope with that danger. Resilience relies on initiative and reciprocity (Woods, 2019). People who are antagonistic to others will not engage in reciprocity and those who are discriminated against will have reduced initiative. Thus, new tensions emerge, and existing ones worsen, fostering more hate speech and other violence inducing behaviors as actors in the community stoke tensions for personal, political, or monetary gain. The spread of these behaviors weakens the community further, and sometimes results directly in violence. This makes it even more difficult for the community as a unit to adapt over time, i.e., degrading the resilience of the community, which provides more and more fuel for violence inducing behaviors.

We have compiled a series of focal topics to direct our future research and recommendations for others. First, we recommend that within community resilience research future work focus on better defining and separating the *Social* and *Community Capital* domains and their indicators. We also recommend that the models highlighted in this review be further analyzed and synthesized into a single comprehensive model. Our project will attempt this synthesis as part of our work. Finally, we recommend work be done to better consolidate research on community resilience, social resilience, societal resilience, community disaster resilience, and all other related threads. Research under each of these titles overlaps considerably and is aligned in their focus on sustained adaptability in

complex systems. Specific inquiry into sustained adaptability and other underlying principles of resilience is currently being carried out by researchers in Resilience Engineering along with several other fields (Woods, 2019; Yu et al., 2020 ; Engle, 2011), which may provide relevant frameworks for such a task. Within online violence prevention and related work, including ours, we recommend that future work directly integrate community resilience models. We will provide a few ideas for how this could be achieved.

Of the models examined in this paper, RESILOC resilience indicators, BRIC, CRIN, the CRI, and CART show the most potential for immediate practical application. The first three (Scherzer et al., 2019; Cutter et al., 2010; Akerkar et al., 2022; City Resilience Index, 2015) provide quantitative measurable proxies for their indicators alongside a comprehensive qualitative structure. These are essentially measures for potential drivers of online violence. Quantifiable measurements could be collected by automated tools and synt hesized in different ways to establish a clear baseline level of community functioning. The resulting information could be used to track the health of a community over time and potentially to prevent the spread of violence online by illuminating driving issues before they boil over. Monitoring only the presence of violent behavior online is inherently reactive. Monitoring the drivers creates more opportunity for prevention and a better means of anticipating future occurrences. For example, it is already well understood that periods of economic difficulty (economic strength is a community resilience indicator) can contribute to anti-migrant sentiment (directly related to multiple kinds of hate speech). Efforts to alleviate economic stress on groups of citizens may prevent the spread of hateful rhetoric by directly addressing the tension it aims to co-opt. These measures also provide a means of evaluating interventions. Because these indicators are agnostic to individual instances of online violence, they can be used repeatedly as the community changes and new sources of tension rise and fall. CRI and CART c an be used to generate complimentary data sourced directly from community members. This data can be used to evaluate the fitness of the other indicators or to assess citizen awareness regarding vulnerabilities in their community. If it is determined that the measurable proxies are insufficient or poorly tailored to a community, the comprehensive qualitative domains of all these models provide a structure that locally specific indicators could fit into. BRIC and CRIN provide a clear example of this. CRIN is an adaptation of the BRIC model to the context of Norway, though the two models share the same structure, their indicators are noticeably different.

CONCLUSION AND FUTURE WORK

Community Resilience is a nascent area of research with the potential to prepare groups of people to better cope with the emergence of threats that transcend national borders. Online violence, as an urgent global issue, presents an opportunity for this research to be tested. The natural place to test this research is prevention work because community resilience models provide means to identify and measure where the underlying architect ure is weak and thus where the tensions on which online violence relies exist in some communities. Because hate speech begets real world violence and degrades the health and wellness of citizens and communities, it fits well into the described domain of interest of community resilience theory. Yet current research has not explored how well community resilience models map to this specific kind of societal disruption. Such an application would serve to strengthen community resilience models by testing them on a novel type of threat and thereby assessing how domain agnostic they are. The insights gained from adapting and revising existing models to address issues of online violence and hate may provide insights and indicators that are similarly useful for dealing with natural disasters and other man-made emergency events.

Future work is needed within community resilience research and highly related fields to synthesize their findings into a single comprehensive framework and in online violence prevention work to integrate community resilience models. We are currently working on synthesizing the models identified in this paper. We will then connect some of the community resilience indicators from our set to different kinds of hate speech. We hope to establish as many similar connections as possible. Once this work is finished, we will develop tools to collect the measurable proxy data for our resilience indicators and to identify multi-modal data as hate speech. Online violence is destructive and dangerous in and of itself and can lead to instances of horrible violence in the real world. It demands that our communities change to overcome it and create a more accepting, diverse, and safe world for all people.

ACKNOWLEDGMENTS

The SOCYTI project has received funding from the Research Council of Norway as a Researcher Project for Technological Convergence related to Enabling Technologies under grant agreement no 331736.

REFERENCES

Berkes, F., & Ross, H. (2013). Community Resilience: Toward an Integrated Approach. *Society & Natural Resources*, *26*(1), 5–20. https://doi.org/10.1080/08941920.2012.736605

City-Resilience-Framework-2015.pdf. (n.d.).

Cutter, S. L., Burton, C. G., & Emrich, C. T. (2010). *Disaster Resilience Indicators for Benchmarking Baseline Conditions*. *7*(1). https://doi.org/10.2202/1547-7355.1732

Eminoğlu, D. N. (n.d.). *INTERCULTURAL DIALOGUE AS A TOOL FOR SOCIAL COHESION*.

Engle, N. L. (2011). Adaptive capacity and its assessment. *Global Environmental Change*, *21*(2), 647–656. https://doi.org/10.1016/j.gloenvcha.2011.01.019

Fitzpatrick, T. (2016). Community Disaster Resilience. In *Disasters and Public Health* (pp. 57–85). Elsevier. https://doi.org/10.1016/B978-0-12-801980-1.00003-9

Gagliardone et al. - 2015—Countering online hate speech.pdf. (n.d.).

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. United Nations Educational, Scientific and Cultural Organization.

Haavik, T. K. (2020). Societal resilience – Clarifying the concept and upscaling the scope. *Safety Science*, *132*, 104964. https://doi.org/10.1016/j.ssci.2020.104964

Hollnagel, E. (2009). The four cornerstones of resilience engineering. *Resilience Engineering Perspectives, Volume 2: Preparation and Restoration*.

Hollnagel, E., Woods, D., & Leveson, N. (2006). Resilience Engineering: Concepts and Precepts. *Resilience Engineering: Concepts and Precepts*.

Jakubowicz, A., Dunn, K., Mason, G., Paradies, Y., Bliuc, A.-M., Bahfen, N., Oboler, A., Atie, R., & Connelly, K. (2017). *Cyber Racism and Community Resilience: Strategies for Combating Online Race Hate*. Springer International Publishing. https://doi.org/10.1007/978 -3-319-64388-5

Keck, M., & Sakdapolrak, P. (2013). What is social resilience? Lessons learned and ways forward. *Erdkunde*, *67*(1), 5–19. https://doi.org/10.3112/erdkunde.2013.01.02

Lupu, Y., Sear, R., Velásquez, N., Leahy, R., Restrepo, N. J., Goldberg, B., & Johnson, N. F. (2023). Offline events and online hate. *PLOS ONE*, *18*(1), e0278511. https://doi.org/10.1371/journal.pone.0278511

Nguyen, H. L., & Akerkar, R. (2020). Modelling, Measuring, and Visualising Community Resilience: A Systematic Review. *Sustainability*, *12*(19), 7896. https://doi.org/10.3390/su12197896

Pfefferbaum, R. L., Pfefferbaum, B., Van Horn, R. L., Klomp, R. W., Norris, F. H., & Reissman, D. B. (2013). The Communities Advancing Resilience Toolkit (CART): An Intervention to Build Community Resilience to Disasters. *Journal of Public Health Management and Practice* , *19*(3), 250–258. https://doi.org/10.1097/PHH.0b013e318268aed8

Pigg, K., Gasteyer, S. P., Martin, K. E., Keating, K., & Apaliyah, G. P. (2013). The Community Capitals Framework: An empirical examination of internal relationships. *Community Development*, *44*(4), 492–502. https://doi.org/10.1080/15575330.2013.814698

RESILOC Resilience Indicators (D3.1 V7.0).pdf. (n.d.).

Scherzer, S., Lujala, P., & Rød, J. K. (2019). A community resilience index for Norway: An adaptation of the Baseline Resilience Indicators for Communities (BRIC). *International Journal of Disaster Risk Reduction*, *36*, 101107. https://doi.org/10.1016/j.ijdrr.2019.101107

The Rabat Plan of Action. (n.d.).

United Nations Strategy and Plan of Action on Hate Speech.pdf. (n.d.).

UNOCT PVE Reference Guide.pdf. (n.d.).

Woods, D. D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. *Reliability Engineering & System Safety*, *141*, 5–9. https://doi.org/10.1016/j.ress.2015.03.018

Woods, D. D. (2019). Essentials of resilience, revisited. In M. Ruth & S. Goessling-Reisemann (Eds.), *Handbook on Resilience of Socio-Technical Systems*. Edward Elgar Publishing. https://doi.org/10.4337/9781786439376.00009

Yu, D. J., Schoon, M. L., Hawes, J. K., Lee, S., Park, J., Rao, P. S. C., Siebeneck, L. K., & Ukkusuri, S. V. (2020). Toward General Principles for Resilience Engineering. *Risk Analysis*, *40*(8), 1509–1537. https://doi.org/10.1111/risa.13494

Scoping Review: Digitalisation in the organisational culture of safety and security agencies

Thomas Kox
Weizenbaum Institute
thomas.kox@weizenbaum-institut.de

Patricia Schütte
Bergische Universität Wuppertal
schuette@uni-wuppertal.de

Sebastian Rohwedder
Weizenbaum Institute
sebastian.rohwedder@weizenbaum-institut.de

Till Büser
Weizenbaum Institute
till.bueser@weizenbaum-institut.de

ABSTRACT

Digitalisation has a significant impact on most aspects of society. State agencies with safety and security tasks have to develop effective strategies to manage the opportunities and risks associated with technological innovations and increased networking. Emerging digital technologies have the potential to affect work practices and thus the organisational culture of such agencies. To examine the extent to which scientific research has addressed the impact of digitalisation on organisational culture so far, a systematic scoping review was conducted. Preliminary results show that the overall research topic is still in an exploratory phase and studies are primarily from the global north. The analyses show that the majority of studies focuses on artefacts around information data leading to respective work practices around data work.

Keywords

Organisational Culture, Digitalisation, Artefacts, Work Practices

INTRODUCTION

Since the advent of digital technologies and computers, digitalization has been a much-discussed development. New information and communication technologies, technical infrastructures and systems, changes in computing power, possibilities and capacities for data collection, processing, analysis and interpretation, artificial intelligence (AI), etc. are combined with significantly expanded corridors of connectivity, networking and exchange in working and living environments. Processes of digital change target all levels: Society and social subsystems, organizations and individuals. In times of the COVID-19 pandemic, these conditions seemed to offer both solutions and options for retreating from the pandemic-influenced world of presence into digital spheres and virtual spaces. In addition to the above-mentioned developments in society as a whole and the relatively dynamic development of technology and services (e.g., IT, AI), the digital transformation has been on the political agenda in several nations for several years and is embedded in a formal framework, for example, in Germany with the constantly updated "Implementation Strategy of the Federal Government". It is therefore no surprise that the field of government agencies including those responsible for public safety and security (in Germany subsumed under the collective term 'authorities and organizations with security tasks' (in German "*Behörden und Organisationen mit Sicherheitsaufgaben*"; BOS) are also under the influence of digital transformation processes (Schütte & Kox, 2022). Recent security-related developments (e.g., cybercrime, digital networking of offender groups), BOS and their representatives, politics, media, and academia (see e.g., Epple & Ludewig, 2019; Hüttelmaier & Käser, 2020) act as 'drivers' with regard to the discussion about the importance of digitalization or the digital transformation of BOS. This is often accompanied by promises or even pledges: Modernity of the organization (and its external image) and achievement of a current 'state of the art', preservation and optimization of the ability of BOS to act under the requirements of digitality and the handling of digital media, (improved) implementation of the mission, various further development potentials, promotion of innovativeness and flexibility, (better and closer) networking

for example in inter-agency cooperation (bitkom, 2020). Over the last years, a lot has happened in the digital transformation of German BOS. A rough insight into different government websites (e.g., of state police forces) contain references to state-specific programmes and master plans for digitalization that clearly show the development stages of the past and coming years. These include the expansion of digital radio and IT infrastructures, the purchase of additional mobile devices (e.g., smartphones for the North Rhine-Westphalian police), the inclusion of digital analysis approaches (such as social media analytics, big data analysis) and the 'purchase' of online specialists who, as 'Intel Officers', provide the police with important real-time information from the social media and other internet sources in particular during ongoing operations (Ministerium Landes Nordrhein-Westfahlen, 2021; Niedersächsisches Ministerium für Inneres und Sport, n.d., 2019). Non-police BOS such as fire departments, aid organizations or the Federal Agency for Technical Relief (THW) are also actively working on their digital transformation. Virtual Operation Support Teams (VOST) have already been established since 2017 as support structures in operational situations and docked with the THW (Fathi et al., 2018). In addition, almost all BOS use social media and operate their own social media accounts (Deutsches Rotes Kreuz, 2018). Innovative efforts are being made to implement new technologies to deal more efficiently with information on future developments in acute operations. Despite these developments, the status of the digital transformation of BOS still leaves room for improvement. It seems that digital transformation is not discussed much together with BOS and infrastructures and equipment are presented as expandable. Overall, there is a certain dissatisfaction with the state of digitalisation in the German BOS with the use of social media as an exception that is now more or less a normal part of communication (blaulicht.digital, 2020). Bureaucratic principles such as written form, file knowledge and rule-boundness offer stable structures for docking digital approaches. However, they may also sometimes slow down decisions, which favour a rather incremental organisational development, which appears to be less compatible with a technology development that brings new offers to the market in a very short time, while previous versions of tools and systems have not yet passed the bureaucratic decision-making processes. The development of operational routines with new technologies takes time. As long as they are not yet a normal part of practice and preparation in calm times, they cannot become a natural tool for acute situation management in turbulent situations, as they increase uncertainty more than they reduce it. The lack of techniques to quickly deal with or practice dealing with new technologies in the organisation delays the utilisation of new technologies (Schütte & Kox, 2022).

OBJECTIVE

Against the background of these considerations, the question arises as to what extent digital technologies are already anchored in the BOS. This question can be theoretically embedded the investigation of organizational culture (e.g., Bellot, 2011; Berg & Wilderom, 2004; Schein, 2010). This approach is not always very clear and free of criticism. Nevertheless, it can be stated that organizational cultures exist, are inherently fuzzy and socially constructed, are a product of groups not individuals, relatively unique to each organisation and subject to continual change (Bellot, 2011).

Schein (2010) assumes that organizational culture emerges at three different levels of an organization. He refers to the top and most visible level as the *artefact level*. This level includes everything that is directly visible, audible and tangible, such as dress code, architecture, office design, technologies, stories and documents, but also behavior patterns. The second level refers to *espoused values*, i.e. the organisations' stated values and rules of behaviour. The values are based on fundamental shared *basic assumptions* that are subconscious or unconscious, barely visible or observable and are almost impossible to research. In addition, van den Berg and Wilderom (2004, p. 571) "*define organisational culture as shared perceptions of organizational work practices*". They see *work practices* as a better means of capturing organisational culture, as these differ significantly more between organisations than values. Adopted from Schein (2010) and van den Berg and Wilderom (2004) we distinguish hereinafter between artefacts and work practices, which are defined as follows:

- *Artefacts*: This category refers exclusively to the field of digital technologies. Artefacts include hardware and software, technical systems, data, means of communication, etc.
- *Work practices*: This refers to activities of organizations and individuals that involve action, execution of processes, individual routines and techniques, etc. related to or shaped by digital technologies.

In order to systematically map the research done in this area, a scoping review (Tricco et al., 2018) was conducted. The following research question was formulated: *How are digital aspects i.e., artefacts and practices included in studies of BOS cultures?*

METHODS

Eligibility Criteria

Against the background of the questions and assumptions outlined above, scientific papers were included in the review process that relate and examine aspects of digitality and levels of organizational culture (e.g., technologies as artefacts, practices, values) in BOS. The papers needed to be: published in scientific journals, peer reviewed, published between the period of 2010 and 2023, social-scientific studies on the topic of organisational culture, written in English, focused on BOS (i.e., police, firefighters, emergency response) as well as focused on artefacts, techniques, practices included in studies of BOS cultures.

Papers were excluded if they did not fit into the conceptual framework of the study: opinion papers, journalistic articles, reviews and comments, technical analytics and discussions (engineering, computer science, etc.), only deal with the specifics of cyber-crime and internet-security, pure presentation of technologies/tools for an abstract addressee in the security field, or papers that discuss organizational change only as part of simulation experiments. Also excluded were studies that deal with organizations that: either cannot be seen as BOS, like actors of pre-hospital care, hospital care and emergency units, or are composed of several organizations and transcend national borders, as is the case with international police organizations, or were 'born' in digital era and are naturally influenced by digital aspects, e.g. Virtual Operation Support Teams (VOSTs). As Cobb et al. (2014, p. 890) point out: "*Empirical research on historical disaster events shows response efforts taking four different organizational forms: established, extending, expanding, and emerging. Emerging organizations are ones that have no pre-existing structure and come together as individuals who begin to connect and coordinate their work*".

Information Sources

To find the relevant academic articles the reviewers searched in the open data bases *Web of Science*, *SCOPUS*, *EBSCO* and initially *Google Scholar* for texts published between January 2010 and August 2023. During the search process, *Google Scholar* was dismissed as a source because of the platform's character limit per search and inconsistency in the results displayed on the first ten pages. On the three remaining platforms, the results were chosen when they featured the words of the search string in the title or the abstract and were published by academic, peer-reviewed journals.

Search

The reviewers formulated the search string in three parts, covering the themes BOS, digitalization and workplace culture respectively. The parts that were connected with AND operators consisted of related words that were connected with OR operators. For the term BOS the reviewers differentiated between police organizations and non-police organizations as described above, resulting in two different search strings. The final search strings for police and non-police are shown in Table 1 using SCOPUS as an example. Please note that the original search also included the terms in German. As there were no additional results, the terms are not listed here for space considerations.

Table 1. SCOPUS Search Strings

Sub-Search	Database	Search String
Police	SCOPUS	police OR "law enforcement" AND "organizational culture" OR "work practice" OR "work routine" OR "value system" OR "behavioral norm*" AND digitali* OR technolog* OR "artificial intelligence" OR "machine learning" OR "virtual reality" OR "augmented reality" OR "social media" OR "ICT"
Non-Police	SCOPUS	"fire brigade" OR "fire department" OR "fire services" OR "firefighter" OR "emergency service" OR "emergency response" AND "organizational culture" OR "work practice" OR "work routine" OR "value system" OR "behavioral norm*" AND digitali* OR technolog* OR "artificial intelligence" OR "machine learning" OR "virtual reality" OR "augmented reality" OR "social media" OR "ICT"

Selection

After removing duplicates, the refined search on *SCOPUS*, *Web of Science* and *EBSCO* concluded in a selection of 94 academic peer-reviewed papers (identification phase) that were then screened for suitability by all four reviewers (see Figure 1).

The first screening each reviewer read all abstracts and marked them as either fitting (n= 27), unfitting (n= 43) or in need of further discussion (n= 24). During this process, inclusion and exclusion criteria were iteratively formulated in more detail to make them clear to all participants. After individual review, the results were contrasted and discussed in the team. The 51 papers that were not unanimously excluded after their discussion were included in the next screening process that consisted of parsing the entire text for suitability. During the collection of said literature, 9 research papers of the 51 research papers were not openly available nor accessible through institutional affiliations. These papers were excluded from the investigation. In similar fashion to the first screening, the remaining 42 papers were read by each reviewer and marked as either fitting or unfitting (screening 2). After reading the entire text, two of the previously fitting papers were excluded and 14 of the papers that still needed discussion were excluded as well. Of the 26 remaining papers, 6 were still in need of further discussion, as the reviewers were in disagreement about their inclusion. After this process three more texts were excluded, leaving n=23 as the total amount of papers to be fully analysed. In the fourth step (inclusion), each of the included full texts were assigned to at least two different reviewers who first read and evaluated each text individually before discussing with the second reviewer.

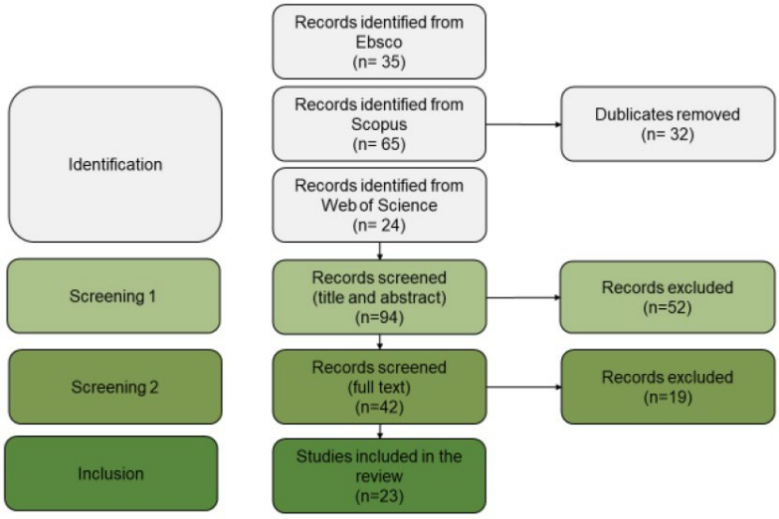


Figure 1. Visualization of the search process with information on excluded texts

Data Charting Process and Data Items

To ensure the most comprehensive data collection possible, an Excel table was developed during the screening steps which all reviewers filled in with categories or variables to be extracted. The categories which are also commonly used in other review procedures are: year of publication, author(s), country, topic (here: police or non-police), focus / aim of the study, study design / methodology, sample / setting / organization type, data collection, results, and limitations. The data in these categories is particularly suitable for text-related, descriptive and quantitative analyses. Some categories such as year of publication, author(s) or keywords are directly extractable from the texts. Other categories such as country or sample / setting can be derived using simple text analysis. The categories artefacts and work practices were derived by the reviewers based on the research question of the scoping review and the theoretical aspects of organization culture. To identify artefact categories, the reviewers assigned a potential artefact label to each text they read. Artefacts were relatively unambiguous in their identification and similar labels were assigned by both reviewers most of the time. In a next step, the reviewers created umbrella terms or overarching categories for all identified artefacts. As work practices are often discussed much less specifically in the texts, a consensus was reached between the reviewers through discussion. This is followed by a more in-depth interpretation of the relationship between the levels in relation to organizational cultural significance and an answer to the research question posed above.

RESULTS

Table 2 lists all 23 reviewed articles together with the listed keywords derived directly from the text. The table provides an initial overview of the different topics, contents and methodological approaches used. Some artefacts and working practices are also already sketched out. By further analysing the texts in depth, we identified 11 of the 23 papers as reporting solely on police organizations, 11 that reported on the various non-police organizations and one paper that focused on both police and non-police organization. A comparison of the geographical distribution of the study areas shows that the majority (seven) of the studies were conducted in the US, five in Sweden and three in the UK. The other study areas include the Netherlands (two), Canada, Italy, Denmark, Germany, Australia and India (one each). No studies that have been included in this review have been published on Africa or South America. In terms of the methods used, 18 studies can be categorised as qualitative approaches (10 of which were ethnographies) and six as quantitative approaches. Three of the 23 papers are a mix of both methods, while two can be described as conceptual papers. Themes, geographical distribution and methods show no noticeable trend over time, with no year having an extraordinary impact on the number of papers published. It should be noted, however, that as one could expect the two conceptual papers are among the first three papers published.

Table 2. List of reviewed papers and listed keywords.

Reference	Keywords
Abrahamson & Goodman-Delahunty, 2014	police · information sharing · impediment · culture · infrastructure
Bergstrand & Landgren, 2011	crisis management · information systems · video communication · information environments · in-situ documentation · situation awareness
Borglund et al., 2012	hybrid environment · IT use · police work · police investigations · prostitution · Synergy-4
Borglund & Nulden, 2012	persona · scenario · design of information technology · emergent use of IT
den Hengst & Staffeleu, 2012	-
Dovigo & Redaelli, 2010	control room · emergency calls · ethnography · ethnomethodology · knowledge management · technology blending
Fischer et al., 2015	tabletop · disaster response · ethnography · collaboration · situation awareness · uncertainty · crisis informatics
Forsgren et al., 2018	policing · social media · activity theory · work activities
Guzik et al., 2021	policing · technology · organisational routines · body worn cameras · socio-materiality
Jensen et al., 2021	ambulances · electronic health record · cultural anthropology · emergency medical services
Landgren & Bergstrand, 2016	situation rooms · coordination · collaboration · work-practice · ethnography · crisis response
McQuade, 2016	-
Piza & Moton, 2023	systematic social observation; police discretion; police decision-making; closed-circuit television (CCTV)
Reuter et al., 2019	head-mounted displays · smartphone · rescue dog units · emergency management · collaboration · empirical study · prototype · evaluation
Robinson et al., 2015	computer supported cooperative work (CSCW) · emergency response · human infrastructure · improvisation
Samuelsson & Berner, 2013	ambulance care · telemedicine technology · ethnography · Sweden · science and technology studies (STS)
Schlosser & Matthews, 2022	CCS concepts· participatory design· design method· secondary video· head-worn display· augmented reality· emergency medical services
St. Denis et al., 2014	crisis informatics· emergency management· public information· social media· trusted volunteers

Tyagi & Dhar, 2017	-
Verhulst & Rutkowski, 2018	group support systems · sociomateriality · affordances · constraints · imbrication · collaborative technology · high reliability organizations
Wilson-Kovacs, 2021	-
Zhang, Joy, et al., 2021	CCS concepts· documentation; data work; electronic health record; teamwork; decision making; distributed cognition; emergency medical services; pre-hospital care; video analysis
Zhang, Sarcevic, et al., 2021	-

In the following we describe the results of the text analysis with regard to artefacts and working practices using the created umbrella terms or overarching categories: In total, nine overarching categories of artefacts have been identified by the reviewers, which are listed in Table 3 alongside the references. These include i) data storage and exchange technologies like cloud storage, ii) data visualisation and analysis technologies such as computer-generated diagrams, interactive maps or figures crafted on whiteboards, iii) information and communications technology (ICT) devices such as cell phones, computers or radio , iv) news and public information sources like news websites or other publicly available information, v) other social and infrastructure data like road traffic data, vi) social media, vii) surveillance technology such as Closed Circuit Television (CCTV) or body worn cameras, viii) various telemedicine technologies including electronic health records, and ix) virtual reality (VR) and augmented reality (AR) technology such as smart glasses or augmented bird tables.

Table 3. Overarching categories of identified artefacts.

Overarching artefact category	References
Data storage and exchange technologies	Guzik et al., 2021
Data visualisation and analysis technologies	den Hengst & Staffeleu, 2012; Fischer et al., 2015; Landgren & Bergstrand, 2016; McQuade, 2016; Tyagi & Dhar, 2017; Verhulst & Rutkowski, 2018
ICT devices	Abrahamson & Goodman-Delahunty, 2014; Bergstrand & Landgren, 2011; Borglund & Nulden, 2012; den Hengst & Staffeleu, 2012; Dovigo & Redaelli, 2010; Verhulst & Rutkowski, 2018
News and public information sources	Landgren & Bergstrand, 2016
Social and infrastructure data	Borglund et al., 2012; Landgren & Bergstrand, 2016
Social media	Bergstrand & Landgren, 2011; Forsgren et al., 2018; Robinson et al., 2015; St. Denis et al., 2014
Surveillance technology	Borglund et al., 2012; Guzik et al., 2021; McQuade, 2016; Piza & Moton, 2023
Telemedicine technologies	Jensen et al., 2021; Samuelsson & Berner, 2013; Zhang, Joy, et al., 2021
VR and AR technology	Fischer et al., 2015; Reuter et al., 2019; Schlosser & Matthews, 2022; Zhang, Sarcevic, et al., 2021

With regard to work practices discussed, in total, six overarching categories of practices have been identified by the reviewers, which are listed in Table 4. These include i) administrative work including resource management and shift planning, ii) community engagement including public communication, iii) data work including information gathering, information processing and information sharing, iv) dispatch and control, v) inter- and intra-agency cooperation and collaboration, and vi) surveillance including investigations in hybrid environments.

Table 4. Overarching categories of identified work practices.

Overarching practice category	References
Administrative work	Borglund & Nulden, 2012; Fischer et al., 2015; Schlosser & Matthews, 2022; Wilson-Kovacs, 2021
Community engagement	Forsgren et al., 2018; Guzik et al., 2021; St. Denis et al., 2014
Data work	Abrahamson & Goodman-Delahunty, 2014; Bergstrand & Landgren, 2011; Borglund et al., 2012; Guzik et al., 2021; Jensen et al., 2021; Landgren & Bergstrand, 2016; Piza & Moton, 2023; Robinson et al., 2015; Samuelsson & Berner, 2013; St. Denis et al., 2014; Tyagi & Dhar, 2017; Verhulst & Rutkowski, 2018; Zhang, Joy, et al., 2021; Zhang, Sarcevic, et al., 2021
Dispatch and control	Dovigo & Redaelli, 2010; Landgren & Bergstrand, 2016
Inter- and intra-agency cooperation and collaboration	Borglund & Nulden, 2012; den Hengst & Staffeleu, 2012; Fischer et al., 2015; Guzik et al., 2021; Jensen et al., 2021; Landgren & Bergstrand, 2016; Reuter et al., 2019; Zhang, Sarcevic, et al., 2021
Surveillance	Guzik et al., 2021; McQuade, 2016; Piza & Moton, 2023

Artefacts around information data could be summarised into a) information-gathering technologies, b) information analysis technologies and c) information visualisation and communication technologies leading to respective practices around data work. When looking at the multitude of work practices researched, data work stands out with more than half of the reviewed literature focusing on this practice. Many of the work practices identified are direct communication processes, either within the agency, between agencies or with/to third parties. There are also work practices that can be categorised as (computer-aided) information processing in the broadest sense, which can be summarised under the umbrella term computer supported cooperative work (CSCW). Finally, it should be emphasised that the categories listed in Table 3 and Table 4 are work in progress and represents preliminary evaluations. There is not always a clear distinction between the categories as they are themselves still under evaluation.

CONCLUSION

After presenting the main steps of the scoping review and addressing some initial findings, this paper draws the following conclusions. We give initial answers to the research question by giving a first descriptive assessment of the literature included in the review and provide an outlook on the next research steps.

The articles examined show that digital aspects have found their way into studies dealing with BOS and their work cultures. The analyses show that the majority of the reviewed studies are qualitative approaches including especially ethnographic studies. This could indicate that the overall research topic is still in an exploratory phase, i.e. there are many blind spots that require an open approach. It is also striking that the reviewed studies are primarily from the global north. The in-depth text analysis show that the majority of studies focuses on artefacts around information data leading to respective work practices around data work.

The team of reviewers assume that digital technologies have arrived in organizations as artefacts, but to a lesser extent in practices, espoused values or even basic assumptions. In other words, in seams that compared to analogue technologies such as pen and paper, axes, guns or water hoses, digital technologies have not yet become fully culturally embedded. Yet, our preliminary analysis has so far only focused on the categories of artefacts and practices, while other categories related to organizational culture have not yet been investigated and are the subject of the next steps.

ACKNOWLEDGMENTS

The contribution of Thomas Kox, Sebastian Rohwedder and Till Büser was part of the joint project "Weizenbaum Institute for the Networked Society" (FKZ 16DII131), funded by the German Federal Ministry of Education and Research.

REFERENCES

Abrahamson, D. E., & Goodman-Delahunty, J. (2014). Impediments to information and knowledge sharing within policing: A study of three canadian policing organizations. SAGE Open, 4(1). https://doi.org/10.1177/2158244013519363

Bellot, J. (2011). Defining and Assessing Organizational Culture. Wiley Periodicals, Inc. Nursing Forum, 46(1), 29–37.

Berg, P. T. van den, & Wilderom, C. P. M. (2004). Defining, Measuring, and Comparing Organisational Cultures. Applied Psychology: An International Review, 53(4), 570–582. https://doi.org/10.1111/j.1464-0597.2004.00189.x

Bergstrand, F., & Landgren, J. (2011). Information systems in crisis. https://doi.org/10.1016/B978-1-84334-647-0.50007-2

bitkom. (2020). Positionspapier. Digitalisierung für die Öffentliche Sicherheit—Beschaffung, Innovation, Funktionalität und Sicherheit ins Gleichgewicht bringen. https://www.bitkom.org/sites/default/files/2020-07/200710_positionspapier_digitalisierung-fur-die-offentliche-sicherheit.pdf.

blaulicht.digital (Ed.). (2020). Digitale Transformation in der zivilen Gefahrenabwehr. https://www.blaulicht.digital/wp-content/uploads/sites/6/2020/09/Studie-Digitale-Transformation-in-der-zivilen-Gefahrenabwehr.pdf

Borglund, E., & Nulden, U. (2012). Personas in Uniform: Police Officers as Users of Information Technology. AIS Transactions on Human-Computer Interaction, 4(2), 92–106. https://doi.org/10.17705/1thci.00041.

Borglund, E., Öberg, L.-M., & Slumpi, T. P. (2012). Success Factors for Police Investigations in a Hybrid Environment: The Jämtland Police Authority Case. International Journal of Police Science and Management, 14(1), 83–93. https://doi.org/10.1350/ijps.2012.14.1.264

Cobb, C., McCarthy, T., Perkins, A., Bharadwaj, A., Comis, J., Do, B., & Starbird, K. (2014). Designing for the deluge: Understanding & supporting the distributed, collaborative work of crisis volunteers. Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW, 888–899. https://doi.org/10.1145/2531602.2531712

den Hengst, & Staffeleu, E. (2012). Different Information Organizations to Produce the Same High Quality Intelligence: An Overview of the Police Forces in the Netherlands1. Policing: A Journal of Policy & Practice, 6(2), 187–193. https://doi.org/10.1093/police/pas005.

Deutsches Rotes Kreuz (Ed.). (2018). Web 2.0 und Soziale Medien im Bevölkerungsschutz. Teil 1: Wissenschaftliche Erkenntnisse über die Nutzung von Sozialen Medien in Krisen- und Katastrophenlagen aus Perspektive des Deutschen Roten Kreuzes.

Dovigo, F., & Redaelli, I. (2010). Knowledge Management in Locating the Patient in an Emergency Medical Service in Italy. COMPUTER SUPPORTED COOPERATIVE WORK-THE JOURNAL OF COLLABORATIVE COMPUTING AND WORK PRACTICES, 19(5), 457–481. https://doi.org/10.1007/s10606-010-9118-7

Epple, G., & Ludewig, F. (2019). Sicherheit im Einsatz durch Open Source Intelligence in Einsatzleitstellen“ (SENTINEL. Polizeispiegel, 20–22.

Fathi, R., Tondorf, V., Schulte, Y., Schütte, P., & Fiedrich, F. (2018). Lageinformationen aus den sozialen Netzwerken: Virtual Operations Support Teams (VOST) international im Einsatz. Notfallvorsorge, 49(2), 4–12.

Fischer, J. E., Reeves, S., Rodden, T., Reece, S., Ramchurn, S. D., & Jones, D. (2015). Building a birds eye view: Collaborative work in disaster response. Conference on Human Factors in Computing Systems - Proceedings, 2015-April, 4103–4112. https://doi.org/10.1145/2702123.2702313

Forsgren, E., Gritt, E., & Allen, D. (2018). Investigating the influence of social media on police activities. International Conference on Information Systems 2018, ICIS 2018. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85062559566&partnerID=40&md5=8fd5e9898d1d4ac109e3c37164256717

Guzik, K., Sesay, A., Oh, O., Ramirez, R., & Tong, T. (2021). Making the material routine: A sociomaterial study of the relationship between police body worn cameras (BWCs) and organisational routines. POLICING & SOCIETY, 31(1, SI), 100–115. https://doi.org/10.1080/10439463.2019.1705823

Hüttelmaier, T., & Käser, B. (2020). Bevölkerungsschutz 4.0. Risiken und Chancen von IT-Systemen im Einsatz. Bevölkerungsschutz, 2, 43–45.

Jensen, F. B., Ladefoged, K. T., Lindskou, T. A., Søvsø, M. B., Christensen, E. F., & Teli, M. (2021). Understanding the effect of electronic prehospital medical records in ambulances: A qualitative observational study in a prehospital setting. International Journal of Environmental Research and Public Health, 18(5), 1–9. https://doi.org/10.3390/ijerph18052330

Landgren, J., & Bergstrand, F. (2016). Work practice in situation rooms – An ethnographic study of emergency response work in governmental organizations. Lecture Notes in Business Information Processing, 265, 157–171. https://doi.org/10.1007/978-3-319-47093-1_14

McQuade, B. I. (2016). Police and the post-9/11 surveillance surge: "Technological dramas" in "the bureaucratic field". Surveillance and Society, 14(1), 1–19. https://doi.org/10.24908/ss.v14i1.5291

Ministerium Landes Nordrhein-Westfahlen. (2021). Digitalisierung. https://polizei.nrw/digitalisierung.

Niedersächsisches Ministerium für Inneres und Sport. (n.d.). Digitale Polizei. https://www.mi.niedersachsen.de/startseite/themen/it_bevollmachtigter_der_landesregierung/masterplan_digitalisierung/digitale-polizei-168213.html.

Niedersächsisches Ministerium für Inneres und Sport. (2019). 13 „Intel Officer" als Verstärkungen für die niedersächsische Polizei. https://www.mi.niedersachsen.de/aktuelles/presse_informationen/13-intel-officer-als-verstaerkungen-fuer-die-niedersaechsische-polizei--174810.html.

Piza, E. L., & Moton, L. N. (2023). Proactive monitoring and operator discretion: A systematic social observation of CCTV control room operations. Journal of Criminal Justice, 86. https://doi.org/10.1016/j.jcrimjus.2023.102071

Reuter, C., Ludwig, T., & Mischur, P. (2019). RescueGlass: Collaborative Applications involving Head-Mounted Displays for Red Cross Rescue Dog Units. COMPUTER SUPPORTED COOPERATIVE WORK-THE JOURNAL OF COLLABORATIVE COMPUTING AND WORK PRACTICES, 28(1–2), 209–246. https://doi.org/10.1007/s10606-018-9339-8

Robinson, J. J., Maddock, J., & Starbird, K. (2015). Examining the role of human and technical infrastructure during emergency response. ISCRAM 2015 Conference Proceedings - 12th International Conference on Information Systems for Crisis Response and Management, 2015-January, 311–320. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84947760505&partnerID=40&md5=cba7772ff45b065cc245f4307f4c2206

Samuelsson, T., & Berner, B. (2013). Swift Transport versus Information Gathering: Telemedicine and New Tensions in the Ambulance Service. Journal of Contemporary Ethnography, 42(6), 722–744. https://doi.org/10.1177/0891241613488941.

Schein, E. H. (2010). Organizational culture and leadership (4th ed.). San Francisco.

Schlosser, P., & Matthews, B. (2022). Designing for Inaccessible Emergency Medical Service Contexts: Development and Evaluation of the Contextual Secondary Video Toolkit. Conference on Human Factors in Computing Systems - Proceedings. https://doi.org/10.1145/3491102.3517538

Schütte, P., & Kox, T. (2022). Vor die Lage – Jetzt und morgen. Herausforderungen von BOS im Umgang mit neuen Technologien und Digitalisierung. In A. Fekete (Ed.), Kritische Infrastruktur und Versorgung der Bevölkerung (pp. 5–25). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-65047-9_2

St. Denis, L. A., Palen, L., & Anderson, K. M. (2014). Mastering social media: An analysis of Jefferson County's communications during the 2013 Colorado floods. ISCRAM 2014 Conference Proceedings - 11th International Conference on Information Systems for Crisis Response and Management, 737–746. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84905841710&partnerID=40&md5=55681700a6280d3fd6882933645e1192

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Annals of Internal Medicine, 169(7), 467–473. https://doi.org/10.7326/M18-0850

Tyagi, A., & Dhar, R. (2017). Knowledge Sharing as a Mediator between Organization Culture and Police Investigation Performance-Moderating Role of Technology: A Study of Tourist Destinations of Uttarakhand, India. POLICING-A JOURNAL OF POLICY AND PRACTICE, 11(2), 185–201. https://doi.org/10.1093/police/paw042

Verhulst, M. J., & Rutkowski, A.-F. (2018). Decision-Making in the Police Work Force: Affordances Explained in Practice. GROUP DECISION AND NEGOTIATION, 27(5, SI), 827–852. https://doi.org/10.1007/s10726-018-9587-5

Wilson-Kovacs, D. (2021). Digital media investigators: Challenges and opportunities in the use of digital forensics in police investigations in England and Wales. Policing, 44(4), 669–682. https://doi.org/10.1108/PIJPSM-02-2021-0019

Zhang, Z., Joy, K., Upadhyayula, P., Ozkaynak, M., Harris, R., & Adelgais, K. (2021). Data Work and Decision Making in Emergency Medical Services: A Distributed Cognition Perspective. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW2). https://doi.org/10.1145/3479500

Zhang, Z., Sarcevic, A., Joy, K., Ozkaynak, M., & Adelgais, K. (2021). User Needs and Challenges in Information Sharing between Pre-Hospital and Hospital Emergency Care Providers. AMIA ... Annual Symposium Proceedings. AMIA Symposium, 2021, 1254–1263.

Human Values for Responsible Decision-Support for Fire Services

Tina Mioch*
University of Applied Sciences Utrecht
Delft University of Technology
tina.mioch@hu.nl

Huib Aldewereld
University of Applied Sciences Utrecht
huib.aldewereld@hu.nl

Mark A. Neerincx
Delft University of Technology
M.A.Neerincx@tudelft.nl

ABSTRACT

Artificial Intelligence systems are more and more being introduced into first response; however, this introduction needs to be done responsibly. While generic claims on what this entails already exist, more details are required to understand the exact nature of responsible application of AI within the first response domain. The context in which AI systems are applied largely determines the ethical, legal, and societal impact and how to deal with this impact responsibly. For that reason, we empirically investigate relevant human values that are affected by the introduction of a specific AI-based Decision Aid (AIDA), a decision support system under development for Fire Services in the Netherlands. We held 10 expert group sessions and discussed the impact of AIDA on different stakeholders. This paper presents the design and implementation of the study and, as we are still in process of analyzing the sessions in detail, summarizes preliminary insights and steps forward.

Keywords

Values, Fire Services, Responsible AI, Decision-Support.

INTRODUCTION

In the last years, AI and data-driven systems are more and more being introduced into first response. Examples are real-time detection, monitoring, and analysis of threats and hazards (e.g., fire detection by means of machine learning algorithms (Talaat & ZainEldin, 2023)), real-time, continuous surveillance of the incident scene by means of drones or ground robots (e.g., by means of a machine-learning-based object detection system (Martinez-Alpiste et al., 2021)), and monitoring and analyzing physiological data of first responders (FRs) to detect (health) threats (e.g., by monitoring heart rate variability to determine current stress levels (Meina et al., 2020)). The development and deployment of AI systems raises new ethical issues (Stahl, 2021) and should be done responsibly, taking ethical, legal, and societal aspects into account. In the last years, many research programs have been started to further research how AI systems can be designed and built responsibly, both by industry (e.g., Google[1] and IBM[2]) and academics (e.g., Harvard[3]); also, law and regulations on responsible AI are planned (e.g., the EU Act (European Commission, 2021)). In addition, several guidelines have been developed for the responsible application of AI systems (AI HLEG, 2019; ECP, 2019). An important principle of all these initiatives is that the context in which AI systems are applied largely determines the ethical, legal, and societal impact of these applications and how to deal with this impact responsibly. To further investigate the responsible application of AI systems for first response, we have started a research project to set up a generic framework of ethical aspects of AI systems for first response. Our research is inspired by the Value Sensitive Design (VSD) methodology, which accounts for human values

*corresponding author
1https://research.google/research-areas/responsible-ai
2https://www.ibm.com/impact/ai-ethics
3https://cyber.harvard.edu/topics/ethics-and-governance-ai

throughout the design process (Friedman & Hendry, 2019). In a previous study (Mioch et al., 2024), first steps were made towards this generic framework by identifying relevant stakeholders and identifying and analysing a first set of human values of these stakeholders regarding AI systems for fire services. These general values have been identified to be important for this domain and should be taken into account when developing AI systems for FR. However, the applicability of the value set depends on the specific AI system and context of use and should be evaluated for each application.

In this research, we further develop and deepen the identified relevant human values for a concrete use case, the development of a specific AI application for the fire services, the AI-based Decision Aid (AIDA). AIDA is a decision-support system that is currently under development by the Dutch Safety Region *Rotterdam-Rijnmond*. The system will structure, analyze, and visualize large amounts of data that are generated (partly in real-time), such as camera pictures of drones or body cams, characteristics of incidents, building properties, historic incident data, etc. This introduction of AI technology offers new possibilities for more efficient and effective decision-making and operations, e.g., prediction of the development of an incident or advice on decisions based on previous incidents; however, the design and development should be done responsibly, taking human values of relevant stakeholders into account.

Research questions in the project concern the identification of expectations of relevant stakeholders, consequences for the organizational structure, how the introduction of AIDA influences work processes and information processing, and which values and value tensions occur when using AIDA in future scenarios. In this research, to identify the impact on human values of relevant stakeholders of the system, we conducted 10 focus group sessions with fire service personnel in which we assess and analyze the stakeholders' values. This article discusses the design and implementation of the study and, as we are still in process of analyzing the sessions in detail, summarizes preliminary insights.

In the following, we first present background on human values for AI development and give an overview of the envisioned AI system AIDA. We then describe the design and implementation of the study and preliminary results, followed by an outlook on next steps and discussion of the results.

BACKGROUND

In this section, we describe the importance of including human values into AI development for FR, give a short overview of previous research on the identification of affected human values, and describe the envisioned decision-support system AIDA, its functionalities and how it will be integrated into the current situation.

Including Human Values into AI Development for FR

In the last years, a lot of attention has gone towards the responsible use and development of AI, amongst others by the EU High-level expert group on AI (AI HLEG, 2019) and IEEE (IEEE, 2021). The AI HLEG identified 4 ethical principles that AI systems should adhere to (i.e., respect for human autonomy, prevention of harm, fairness, and explicability) and 7 key requirements (i.e., human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability). The AI HLEG stresses that these ethical principles and requirements need to be situated and considered in context and that it should be determined for the specific applications whether these requirements are applicable, possibly supplementing them with additional relevant requirements. This means that for all AI applications, the ethical aspects need to be investigated to determine relevancy and applicability. IEEE also emphasizes the importance of considering ethical values throughout the stages of concept exploration and development.

The importance of including human values into the design and development of (AI) systems is also at the basis of the Value-sensitive design methodology (VSD) (Friedman et al., 2013). VSD is an iterative methodology that integrates three perspectives, namely conceptual, empirical, and technical investigations. These investigations are executed iteratively, though not necessarily in a set order. In this research, we focus on the empirical investigations for identifying impact of AIDA on fire services.

In previous work, ethical issues have been investigated for example for Search and Rescue robotics. Harbers et al. (Harbers et al., 2017) identified and analyzed relevant values affected by rescue robotics; values that were identified were, amongst others, personal safety, safety of others, access to information, well-being, effectiveness, ease of use, authority, health, and contact. Battistuzzi et al. (Battistuzzi et al., 2021) conducted a scoping review on ethical concerns in rescue robotics, identifying seven core ethically relevant themes: fairness and discrimination; false or excessive expectations; labor replacement; privacy; responsibility; safety; trust. In Mioch et al. (2024), as a

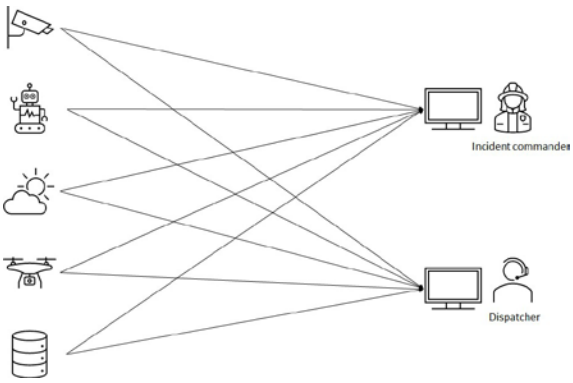


Figure 1. Current system set-up.

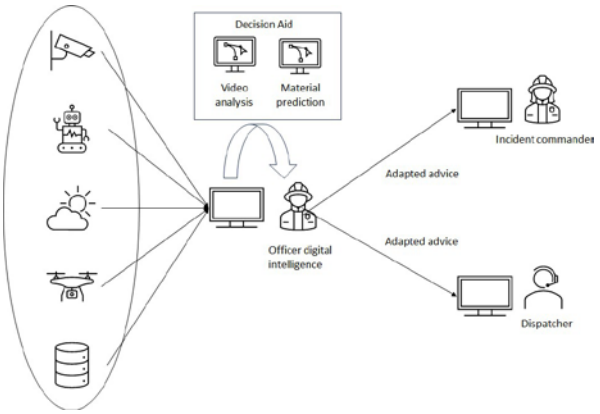


Figure 2. Envisioned set-up with AIDA.

step towards a generic framework of ethical aspects for AI in FR, relevant direct and indirect stakeholders of AI systems for fire services were identified and relevant human values of a prioritized selection of these stakeholders regarding the application of AI systems were investigated. The following stakeholders were identified as most relevant to take into account during the introduction of AI systems: fire fighters, incident commanders, special operations fire fighters (e.g., for hazardous materials and digital exploration), dispatchers, company doctors, incident researchers, citizens, and the surroundings (being an umbrella term for the nearby environment, including objects, people, nature, etc.). Positive as well as negative impact was identified for different human values. Starting point of the identification of values was the account of Friedman et al. (Friedman et al., 2017) as these are considered to be particularly important for technological design. Several values as defined by Friedman et al. were found to be impacted (mostly) positively for the different stakeholders (although there was no value that is only affected positively), i.e., *physical well-being* and *psychological well-being*. Several values were found to be impacted only negatively for the different stakeholders, i.e., *autonomy*, *identity*, *informed consent*, *privacy*, and *trust*. Several other values were identified to be impacted by AI systems, i.e., *accountability*, *transparency*, *reliability*, *security*, and *appropriate training*. AI systems should (explicitly) support the possible positive impact that they could have whereas the negative impact should either be reduced or avoided altogether, e.g., by specifying design requirements that explicitly take these human values into account. In this research, we will further investigate the impact on human values of the (envisioned) AI system AIDA, situating and applying the research by Mioch et al. (2024) for this specific application to further determine relevance of found values.

AIDA

A lot of different data is being generated or accessible by fire services, such as camera images of drones or body cams and building properties; currently, this data is streamed directly to FRs without any analysis or interpretation and displayed on different (sub)screens, see Figure 1. That means that FRs continuously need to monitor large amounts of data to build good situation awareness (SA). This is very time– and work-intensive and can lead to cognitive overload of individual FRs and subsequently to suboptimal decision-making, especially during complex and dynamic incidents and high time pressure. To mitigate these effects, the Dutch Safety Region *Rotterdam-Rijnmond* is developing an AI system, AIDA, to support them in the analysis, interpretation, and visualization of this (real-time) data. This introduction of AI technology offers new possibilities for more efficient and effective decision-making and operations, e.g., prediction of the development of an incident based on data from comparable incidents and information on the current status. In addition, new roles are envisioned to support in the sense-making process and sharing of information.

For AIDA, different modules are being developed. For this research, we will focus on two modules that are designed to help the dispatcher and the incident commander with building up situation awareness and their decision-making, i.e., the *video analysis module* and the *material prediction module*. The main goal of the former is to support shared situation awareness, whereas the latter will provide advice on the material to send to an incident. For both modules, AI systems are envisioned. Shared situation awareness (of for example the incident commander) in the *video analysis module* will be supported by analysing and interpreting different video sources (e.g., drone camera pictures) to identify relevant information, such as locating humans in camera or drone pictures in the incident area faster and with less cognitive effort, aiding with decision-making regarding the deployment of teams and logistical effort (a first prototype description can be found in Heemskerk et al. (2024)). This will be particularly useful in situations in which the situation is very dynamic and complex. Advice in the *material prediction module* will

Table 1. Overview over scenario selection per session and participants.

Session	*Scenario*	*# (Head) Incident Commanders*	*# Dispatchers*	*# Fire fighters*	*# Emergency co-ordinators*
1	(Adapted) Scenario 2	2	2		
2	Scenario 2	1	1		1
3	Scenario 1	2	1		
4	Scenario 2	2	2	1	
5	Scenario 1	1	1		
6	Scenario 1	2	2		
7	Scenario 2	2	2	1	
8	Scenario 1	2	2		
9	Scenario 2	1	3	1	
10	Scenario 1	2	3		
Total		*17*	*19*	*3*	*1*

be given on, amongst others, the expected material that is needed during particular incidents, based on data of previous incidents, real-time characteristics of incidents and other data such as cadastre data. The safety region expects to introduce a new role, the *Officer Digital Intelligence (ODI)*, who monitors and analyses the different data streams and outputs of the different modules and coordinates the information towards the dispatcher and incident commander. Both keep access to all data, but are supported by the ODI, see Figure 2. In the following, the two modules are described in more detail.

Video Analysis

A lot of camera pictures are generated by different cameras during incidents. One source of video data are drones that film the location of an incident from above. The video feeds need to be monitored to recognize relevant information for incident response, such as development of the fire or presence of people. The envisioned AI module will analyze and interpret the video stream and detect humans. It will be able to classify humans into fire fighters, ambulance personnel, and civilians. In addition, a short summary can be made of the video stream based on changes and development of presence of people in the footage.

Material Prediction

A lot of data is available from past incidents, as the development of each incident is logged, e.g., the address of an incident, characteristics of the incident, and material that was requested (type of vehicle, specialization of personnel, etc.). The envisioned AI system will, based on correlations found in the available past data, predict what kind of material is needed for an incident and advise the officer to request the necessary material.

METHOD

To determine the relevant values of different stakeholders for AIDA, we conducted 10 focus group sessions with different stakeholders of the fire services. In total, 40 employees of the fire services participated, of which 1 head incident commander, 16 incident commanders, 19 dispatchers, 3 fire fighters, and 1 emergency coordinator. Many of the participants had several roles in the fire services, e.g., a dispatcher also working as voluntary fire fighter; for this overview, we selected the role that the participants mentioned when signing up. 5 participants were female. For an overview of the participants, see Table 1.

During the focus group sessions, the following steps were followed: first, a short introduction into AI technology was given, its working, possibilities, and limitations, to make sure that participants had a general knowledge of the working of AI systems to be able to identify impact of AI. Most participants had heard of the plans to introduce AIDA, but were not familiar with details. For that reason, the envisioned functionalities of AIDA were introduced as described above, together with the envisioned embedding into the organisation. Two scenarios, one for each module, were used to instantiate the working of AIDA into situation and organizational processes, see Table 2. These scenarios highlighted value tensions. In each session, one of these scenarios was presented as a basis for discussion on the positive and negative impact of the corresponding AIDA module. As a first step for this research, the participants were asked to identify positive and negative impact of AIDA for the roles *Incident commander*, *Dispatcher*, and *Officer Digital Intelligence*. These roles were chosen because (most) participants belonged to at least one of these stakeholder groups, these stakeholder groups will be directly influenced by the results of the

Table 2. Scenario descriptions.

	Scenario 1 - Video analysis module	*Scenario 2 - Material prediction module*
Situation	Large fire in industrial building; The fire brigade explores the situation. A drone team is called and a drone provides a live camera feed of the building and environment. It becomes clear that not all employees are accounted for.	A building in the center of Rotterdam is on fire. The dispatcher receives the 112 call with the location of the fire and notifies the fire brigade with the information; according to procedure, two cars are sent.
Data	Drone camera stream; camera pictures of bodycams.	Historical incident data, characteristics of current incident.
AI technology	The AI system processes the data in real-time and combines the various data sources; it detects humans in the camera pictures and distinguishes between fire fighters, ambulance personnel, and civilians. The system can also give a summary of the video feed.	The AI system predicts, based on historical incident data, what kind of material will be needed. In an adapted version (session 1), the AI system predicts the personnel that should be called up to be deployed (instead of which material should be deployed).
Steps	The ODI requests a summary of the last minutes of video stream, in which newly recognized persons are displayed in the video stream. AIDA identifies people on the video stream and shows a summary of all people found. The system distinguishes between firefighters, ambulance personnel, and civilians. The ODI looks at the summary of the video stream. Several citizens are shown. The ODI provides the relevant information to the dispatcher.	AIDA continuously monitors and assesses the material requirements of incidents, based on various data sources such as land registry data and historical incident data. It predicts that only one car is needed for this incident and passes this on to the ODI. The ODI assesses the advice, agrees with it and passes it on to the operator.

system, and because these stakeholder groups are amongst the stakeholders that were identified in previous research as particularly relevant (Mioch et al., 2024). In the continuation of this research, impact on other stakeholders will also be investigated. The results were shared plenary and discussed. During the discussion, additional impact was added. Also, the participants were asked to discuss how positive impact could be supported and negative impact be reduced. During the focus group sessions, we tried to focus on the impact of the AI modules and not on the impact or the introduction of the new role (the *ODI*).

PRELIMINARY RESULTS

As we are still in the process of analyzing the expert group sessions in detail, in this section, we will describe preliminary results, i.e., a selection of expected impact that was mentioned by the participants. First, we shortly describe impact and requirements of the new role that is suggested to support dispatcher and incident commander with the analysis of the data. After, for each of the two AI modules, we describe a selection of impact participants mentioned and corresponding human values that are implicated, based on the previously identified list of human values that are implicated by AI systems for FR (Mioch et al., 2024). As we are still in the process of analysing the results, these first impressions will later be updated.

New Role *Officer Digital Intelligence* As mentioned above, the introduction of AIDA is accompanied with the introduction of a new role, the *ODI*. Various aspects were mentioned in the sessions that are relevant to the role of ODI, and especially to the collaboration with the incident officer and the dispatcher. First of all, the expertise of the ODI is very important; the ODI needs to have good insight into the tasks and role of both dispatcher and incident commander. Since the dispatcher and officer can be very busy with their own tasks, it is very important that the ODI understands what kind of information the other two need or do not need at any time, so as not to disrupt their process, introduce overload, or miscommunication (by speaking 'two languages'). The ODI should take work off the hands of the dispatcher, but should not make decisions; the ODI should have a support function and work together with the dispatcher and the incident commander.

Video Analysis

The participants discussed the impact of the *Video analysis module* in 5 sessions. In the following, we present a selection of expected impact that was mentioned by the participants, sorted by value or ethical aspect that is affected. Please note that this overview is by no means complete yet and only gives an idea on some of the topics that came up.

Autonomy Autonomy refers to people's ability to decide, plan, and act in ways that they believe will help them to achieve their goals (Friedman et al., 2013). Impact that affects *autonomy* was mentioned for the introduction of both modules. Regarding the module *Video analysis*, participants stated that the systems might lead to less decision-making power because of valuing and interpreting information becomes more difficult, e.g., because of not understanding the information and analysis that the system provides ("But the problem with the system is that it might be better than we are, and that we cannot understand it. [...] It is very difficult to find a good balance: to some extent you want to use it, but otherwise you also want to handle the situation yourself."). Participants mentioned that they are concerned about their dependence on the technology and that if the system fails, their ability to operate might be impaired because they forgot how to handle the situation without the system. Specifically for the application of detecting and summarizing persons in the drone video stream, participants were concerned that it would influence autonomy negatively through disproportional trust in the results of the system ("It's exactly like that video where the monkey walks through the screen where 100 people don't see that monkey, [...], you no longer see the rest of the information, because I become triggered to look at people.")

Identity Identity refers to people's understanding of who they are over time, embracing both continuity and discontinuity over time (Friedman et al., 2013). Impact that related to the value *Identity* was mentioned in several sessions, for example regarding expected changes in tasks for the different stakeholders, e.g., positive impact on the dispatcher ("the dispatcher can focus on his or her main tasks"), but also the expectation that some tasks will no longer be performed by the dispatcher ("The ODI takes quite a bit of work off the dispatcher's hands"). Some participants were afraid that their role will become less relevant and that they will be skipped in the communication between ODI and incident commander. The dispatcher's expertise could be used less which led to feeling threatened in their work ("That can [...] be kind of demotivating for the dispatcher or feel as if they become redundant or irrelevant, or less important").

Psychological well-being Psychological well-being falls under the value *Welfare* as identified by Friedman et al. (2013). Participants mentioned that introducing (summaries) of video streams of incidents for the ODI or the dispatcher will give them more information and overview on what is happening and might improve decision-making. However, it will also introduce pictures of what is happening, which might lead to more PTSD (" you see something that you may not want to see in that role").

Trust in AI Participants mentioned that (appropriate) trust in AI systems is important. Participants said that there is a danger of too much trust, e.g., being influenced in decision-making too much, not evaluating and reflecting enough the situation ("I'm a bit afraid of tunnel vision; that we will focus 100 percent on the system [...] and that common sense will fade into the background"). In addition, it is very important to understand the capabilities and limitations of the system to be able to determine how to include the pictures in the decision-making process ("[...] that you are going to make assumptions, that you do not check whether the pictures from the drones are actually current reality, or that there are aspects that are not filmed.").

Accountability Accountability can be seen in the context of liability (legal responsibility) as well as social responsibility. Participants mentioned that they expect negative impact on accountability, for example when the system provides information that is not followed on by the FRs ("And I think you will soon have another problem, that the information is there, [...] but you have not done anything with it. And then you get public pressure [...] and you will get blamed for it.").

Material Prediction

The participant discussed the impact of the *Material prediction module* in 4 sessions. In the following, we present a selection of expected impact that was mentioned by the participants, sorted by value or ethical aspect that is impacted.

Autonomy Regarding *Autonomy*, participants mentioned that it might be more difficult for dispatchers to make decisions, as it might be difficult to interpret the advice and integrate it in the decision-making process ("We now always say the caller is always right, until proven otherwise. So if the caller says there is a raging fire and the system says that based on that criteria, you have to scale up or down, how do I deal with that?"). In addition, they were concerned about depending too much on the system and about their ability to operate when the system fails ("Are you still sharp at the moment the system does something crazy?").

Trust in AI Participants mentioned that they have difficulty in trusting an AI system that gives advice on aspects that might lead to a less optimal task performance, for example, in the case of sending less material to an incident ("I feel anxious (apprehensive) about scaling down based on technology."). In addition, even if they could understand the advice and rationally, this advice seems sound, then these kind of decisions still also have an emotional aspect that makes it difficult to follow the advice ("You may be able to explain it all rationally, but your feeling... says something different."). Also, they mentioned that there always is a difference between a theoretical reality (e.g., houses that are optimally built regarding fire safety, but where the context such as changes in the building or the inhabitants counteract the optimal circumstances) and practice, which makes systems based on data inherently unreliable.

Accountability Participants are concerned that legal responsibility need to be clearly specified; for example, when following advice of the AI system although it is not correct ("If I scale down and it is nevertheless serious, then it's my fault"). In addition, participants feel a strong social responsibility and see it critically that the *Material prediction* could advise less material to be dispatched ("It goes against your sense of helping").

Privacy Privacy refers to a "claim, an entitlement, or a right of an individual to determine what information about himself or herself can be communicated to others" (Friedman et al., 2013). Impact on the value *privacy* was only mentioned during the first session, in which the scenario (and AI module) was adapted towards planning of which personnel should be called up to be deployed (instead of which material should be deployed). During this session, the topic was discussed fervently, as this directly had links to the monitoring of personnel. Impact on privacy for citizens was not mentioned during the sessions. This might be because the participants were not explicitly asked to identify impact for these stakeholders.

DISCUSSION

Previous research (Mioch et al., 2024) investigated which human values are impacted by AI systems for fire services in general; the investigation was set-up to identify these values broadly and the impact of AI systems was not contextualized. This resulted in an overview of values that might be impacted (positively and negatively), but not in concrete (design) requirements on how to support positive impact and decrease negative impact. With this research, we contextualize this previous research by choosing a concrete AI application and discussing the impact of two example modules on fire services practice. By eliciting expected impact for these concrete AI modules, translating this impact towards affected human values and linking design requirements to these values, we hope to be able to make the step towards defining grounded design requirements for the responsible design and development of these modules.

In this paper, we have made first steps towards identifying relevant human values for two modules of AIDA. To achieve this, we have held 10 focus group sessions, 5 for each module, in which we discussed impact of the modules on tasks and roles of the FRs; we take a socio-technical systems approach, expecting changed task allocations and -responsibility and new human-AI dependencies. We presented first results of these sessions, i.e., statements that can be linked to value considerations, which again can influence design requirements. These design requirements should (explicitly) support the possible positive impact on relevant values whereas negative impact should either be reduced or avoided altogether, e.g., by specifying design requirements that explicitly take these human values into account. There are several things that can be noted on the basis of the preliminary results: first of all, several of the previously identified human values are implicated by the AI modules, e.g., *Autonomy* and *Identity*. In addition, values such as *Trust in AI* and *Accountability* have been mentioned for both modules. During the focus group sessions, the participants discussed the impact of the two AI modules within the context of two concrete scenarios and their fire services practice, including obstacles they see for the introduction of these specific modules and possible implementation choices to enhance positive impact; this discussion will be the basis for the analysis of the design requirements.

During the sessions, it was difficult to keep discussion to the specific scenarios and impact of the selected AI module. Possibly, the scenarios we presented were not specific enough and, in the future, should be made more specific by a simulated collaboration environment (as used in e.g., (Beuker et al., 2021)). Also, possibly because of the very different (technical) expertise and background of the participants, the discussions not always focused on the expected AI-related impact, but also on having access to particular data or other impact of IT-related systems currently in use.

We realize that, as described above, the introduction of AIDA is very complex and that it is not (yet) possible to discuss impact on the tasks, processes, and responsibilities of the different roles, as this also includes organisational changes through the introduction of a new role, the *ODI*. We only superficially touched on this issue when describing preliminary requirements on the expertise of the ODI. In the future, we will also look into these organizational aspects in more detail and analyse, discuss, and advise on the organizational and procedural changes. We also realize that for the participants, it was a challenge to differentiate between the impact of AIDA and the impact of the organizational change. We will take this into account in the analysis of the results.

The focus group sessions have been held right at the beginning of the modelling and development process for AIDA. The functionalities of the system are not yet set. The modelling and development process will profit from our research in several ways: (1) based on positive and negative impact that has been identified, design requirements will be specified that support the positive impact and reduce negative impact, (2) explicitly, values of different stakeholders are taken into account during the design and development process, and (3) the different stakeholders will be regularly involved in the research, which facilitates support and acceptance and improves the quality and suitability of the system. For now, we focused this effort of identifying impact of the two AIDA modules on three stakeholder groups from the fire services (i.e., *dispatcher*, *incident commander*, and *officer digital intelligence*); in the future, we would also like to hold focus group sessions with other FR organizations (e.g., police, ambulance services) and citizens to also take their perspectives into account.

In the future, we will further analyse the results of the expert group sessions and plan to (1) compare value implications between the two modules ('Are different values affected by the two modules?') and (2) compare identified values with the general framework ('Are there additional values affected that have not been identified before, or are some previously identified values not affected by these two modules?'). We will then integrate the results into the ethical framework, together with values found from normative sources, adding results of a scoping literature review on ethical aspects of AI systems applied to first response. The ultimate goal is to develop a framework for responsible development and use of AI in the fire services domain, including requirements and methods (e.g., with regard to identifying values and value tensions), models (e.g., with regard to situated values), design patterns of responsible AI development and deployment, and evaluation methods for responsible AI so that practitioners and AI developers can use the framework as a toolbox towards responsible AI design, development, and deployment.

ACKNOWLEDGMENTS

We would like to thank the first responders of *Gezamenlijke Brandweer Rotterdam*, *Veiligheidsregio Rotterdam-Rijnmond*, and *Veiligheidsregio Zuid-Holland-Zuid* for their participation in this research. This research was supported by the University of Applied Sciences Utrecht (HU) through a '*promotievoucher*'.

REFERENCES

AI HLEG. (2019). Ethics Guidelines for Trustworthy AI. European Commission.

Battistuzzi, L., Recchiuto, C., & Sgorbissa, A. (2021). Ethical concerns in rescue robotics: A scoping review. *Ethics and Information Technology*, *23*(4), 863–875. https://doi.org/10.1007/s10676-021-09603-0

Beuker, T., Mioch, T., & Neerincx, M. A. (2021). Team Design Patterns for participatory development of First Response Human-Agent Teaming. In W. Karwowski, T. Ahram, M. Milicevic, D. Etinger, & K. Zubrinic (Eds.), *Human systems engineering and design (IHSED 2021): Future trends and applications. AHFE (2021) international conference* (Vol. 21). AHFE Open Access. https://doi.org/10.54941/ahfe1001151

ECP. (2019, January). *Artificial Intelligence Impact Assessment (English version)*. Retrieved February 5, 2024, from https://ecp.nl/publicatie/artificial-intelligence-impact-assessment-english-version/

European Commission. (2021). *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts.* (final No. COM/2021/206).

Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. Mit Press.

Friedman, B., Hendry, D. G., & Borning, A. (2017). A Survey of Value Sensitive Design Methods. *Foundations and Trends® in Human–Computer Interaction*, *11*(2), 63–125. https://doi.org/10.1561/1100000015

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2013). Value Sensitive Design and Information Systems. In N. Doorn, D. Schuurbiers, I. van de Poel, & M. E. Gorman (Eds.), *Early engagement and new technologies: Opening up the laboratory* (pp. 55–95). Springer Netherlands. https://doi.org/10.1007/978-94-007-7844-3_4

Harbers, M., de Greeff, J., Kruijff-Korbayová, I., Neerincx, M., & Hindriks, K. (2017). Exploring the ethical landscape of robot-assisted Search and Rescue. *Intelligent Systems, Control and Automation: Science and Engineering*, *84*, 93–107. https://doi.org/10.1007/978-3-319-46667-5_7

Heemskerk, J., Mioch, T., Maathuis, H., & Aldewereld, H. (2024). Long-range human detection for first response decision support. In B. Penkert, B. Hellingrath, M. Rode, A. Widera, M. Middelhoff, K. Boersma, & M. Kalthöner (Eds.), *Proceedings of the 21st ISCRAM conference*.

IEEE. (2021). IEEE standard model process for addressing ethical concerns during system design. *IEEE Std 7000-2021*, 1–82. https://doi.org/10.1109/IEEESTD.2021.9536679

Martinez-Alpiste, I., Golcarenarenji, G., Wang, Q., & Alcaraz-Calero, J. M. (2021). Search and rescue operation using uavs: A case study. *Expert Systems with Applications*, *178*, 114937.

Meina, M., Ratajczak, E., Sadowska, M., Rykaczewski, K., Dreszer, J., Balaj, B., Biedugnis, S., Węgrzyński, W., & Krasuski, A. (2020). Heart rate variability and accelerometry as classification tools for monitoring perceived stress levels—a pilot study on firefighters. *Sensors*, *20*(10). https://doi.org/10.3390/s20102834

Mioch, T., Aldewereld, H., & Neerincx, M. A. (2024). Empirical investigation of values affected by AI systems for fire services. In B. Penkert, B. Hellingrath, M. Rode, A. Widera, M. Middelhoff, K. Boersma, & M. Kalthöner (Eds.), *Proceedings of the 21st ISCRAM conference*.

Stahl, B. C. (2021). Ethical Issues of AI. In B. C. Stahl (Ed.), *Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies* (pp. 35–53). Springer International Publishing. https://doi.org/10.1007/978-3-030-69978-9_4

Talaat, F. M., & ZainEldin, H. (2023). An improved fire detection approach based on yolo-v8 for smart cities. *Neural Computing and Applications*, *35*(28), 20939–20954.

Inter-professional simulator training for reliable ship-helicopter interactions

Thomas Lübcke
German Maritime Search and Rescue Service – DGzRS
luebcke@seenotretter.de

Hendrik Wilhelm
Witten/Herdecke University
hendrik.wilhelm@uni-wh.de

Oda Schliebusch-Jacob
German Maritime Search and Rescue Service – DGzRS
schliebusch-jacob@seenotretter.de

Tim Kreter
Witten/Herdecke University
tim.kreter@uni-wh.de

Tanja Martini
German Aerospace Center – DLR
tanja.martini@dlr.de

Torsten Gerlach
German Aerospace Center – DLR
torsten.gerlach@dlr.de

ABSTRACT

This work-in-progress paper presents the preliminary results of an exploratory study on the effectiveness of a simulator training for ad hoc collaboration between sea- and airborne professionals in the field of maritime search and rescue. To enable this training, two different types of simulators were coupled. This coupled simulation environment provided the 49 (preliminary) study participants—seafaring and helicopter professionals—with authentic simulator training conditions. Our training was developed and tested using a two-phase approach. In the first phase, we used qualitative methods to identify requirements and design a training using insights from domain experts. In the second phase—the preliminary results of which are presented here—we tested the impact of the training on the acquisition and transfer of competencies as rated by the participants. To uncover contingencies of these training effects, we explore demographic predictors. We show which set of participants perceived that they benefited most from the training.

Keywords

ad hoc teams, learning, maritime search and rescue, simulation, training

INTRODUCTION

In the event of major incidents, such as natural disasters or major accidents, cross-organizational cooperation—involving, e.g., fire departments, police, rescue services or military—is an effective way to mobilize the necessary resources in a short period of time to save human lives (Hällgren, Rouleau, & de Rond, 2018; Steigenberger, 2016). This requires people from different professions and backgrounds (full-employed rescue workers and volunteers) to solve tasks as an ad hoc team—that is, individuals who collaborate to accomplish a task but do not

know each other, and have no prospect of further collaboration (Hollenbeck, Beersma, & Schouten, 2012; Salas, DiazGranados, Klein, Burke, Stagl, Goodwin, & Halpin, 2008)—often under critical conditions (Flin, O'Connor, & Crichton, 2007; Weick & Sutcliffe, 2015; Weick, Sutcliffe, & Obstfeld, 1999; Weick & Roberts, 1993). The organizations that commission these individuals strive for reliability in fulfilling their (sovereign) mission, and thus train these individuals in a standardized manner (International Maritime Organization [IMO], 2014; IMO 2015). Therefore, clearly defined hierarchies, training programs, and internal career paths are characteristic of these organizations. In addition, regular internal training and live exercises ensure that each individual and the organization as a whole are always able to act. In these settings, learning takes place primarily through practical action by systematically reflecting on operations, e.g., in debriefings or by means of after-action reviews (Klein, 2015; Kolb, 2014). These trainings are known to benefit individual's expertise. However, these trainings rarely cover cross-organizational work in ad hoc teams—where different professions from more than one organization are required to collaborate in a critical situation. In consequence, we know little about how rescue workers can be prepared for such incidents. This presents a problem, because large-scale incidents are projected to become more likely as a consequence of climate change (Birkmann, Liwenga, Pandey, Boyd, Djalante, Gemenne, Leal Filho, Pinho, Stringer, & Wrathall, 2022).

This article addresses this relevant shortcoming in previous research. We provide a preview of an ongoing—soon-to-be-completed—study, which analyzes data from a new cross-professional simulator training course (AMARIS[1] training) linking professionals from the shipping industry (shipmasters, nautical officers, and coxswains) and helicopter emergency medical service (HEMS) personnel, i.e. helicopter crews including pilots, medical crew members, and hoist operators. In this study, we seek to explain the AMARIS-induced competence acquisition—i.e., the improvement of professionals' ability to successfully perform cross-organizational work in ad hoc teams—and competence transfer—i.e., the ability to appropriately apply the acquired knowledge in the workplace—of training participants (Blanchard & Thacker, 2023; Shippmann et al. 2000; Thielsch & Hadzihalilovic, 2020). We do this in order to gain an understanding of whether joint simulator training for members of different professions has the potential to prepare them for future ad hoc team collaboration. To maximize practical relevance and impact, we conducted our study in the field of maritime search and rescue (SAR), a setting where cross-professional collaboration is both common in practice but non-existent in training programs. Our study covers three elements: two data collection phases (qualitative and quantitative) and the development of a technical simulator innovation allowing us to train and test inter-professional ad hoc teamwork. Specifically, we coupled[2] the SAR simulator center of the German Maritime Search and Rescue Service (DGzRS) in Bremen with the Air Vehicle Simulator (AVES) of the German Aerospace Center (DLR) in Braunschweig—two fundamentally different types of simulators that serve as a networked training and laboratory environment for the subsequent data collection. AVES consists of three simulators with a Roll-on/Roll-off (RO-RO) cockpit exchange concept and is in operation since 2013. A full motion and two fixed based simulators with large projection systems can carry cockpit modules and interchange them within a few hours. Currently four cockpit infrastructures are available: an Airbus A320, Airbus Helicopter EC135, Advanced Future Cabin (windowless cabin) and Falcon 2000 ISTAR (DLR, 2024). The SAR simulator is based on the networking of five ship handling simulators, two of which can be used multifunctionally: either as a ship, helicopter (tactical) or rescue control center workstations. The system is completed by a Maritime Rescue Coordination Centre (MRCC) cubicle with four workstations including a coastal radio station workstation and two Mobile Operation Stations (MOS), which can be integrated into simulations via Virtual Private Network (VPN) and are also designed to be multifunctional. All simulator cabins are equipped with communication media (e.g., radios, telephone, and e-mail). A central instructor center operates the SAR simulations. While the SAR simulator is used exclusively for training emergency personnel, the AVES is used purely as a research simulator for the development of information and control systems.

There are several previous examples of simulators being networked for research and training purposes. These include the coupling of a ship handling simulator and an engine room simulator at nautical universities or the transnational linking of ship handling simulators in the European Maritime Simulator Network (EMSN) to research safe vessel traffic management (e.g. STM, 2024). Coupling the AVES with the SAR simulator is innovative in its approach, as it is the first to connect simulation environments across organizational boundaries

[1] AMARIS is an acronym for the German project title "Aeronautische und maritime Innovationsumgebung für interorganisationale Simulationen". Information on the project is available from the authors upon request.

[2] AMARIS required a distributed simulation setup across the internet, because the simulators are located in different facilities. While the AVES uses a software environment developed entirely by the DLR, the industrial standard "Distributed Interactive Simulation (DIS)" was used to interconnect both simulators. The SAR simulator works as a host and is capable of reading and writing the DIS-protocol standard. To ensure safety and security during data exchange, a VPN tunnel was used.

and different domains (Figure 1).

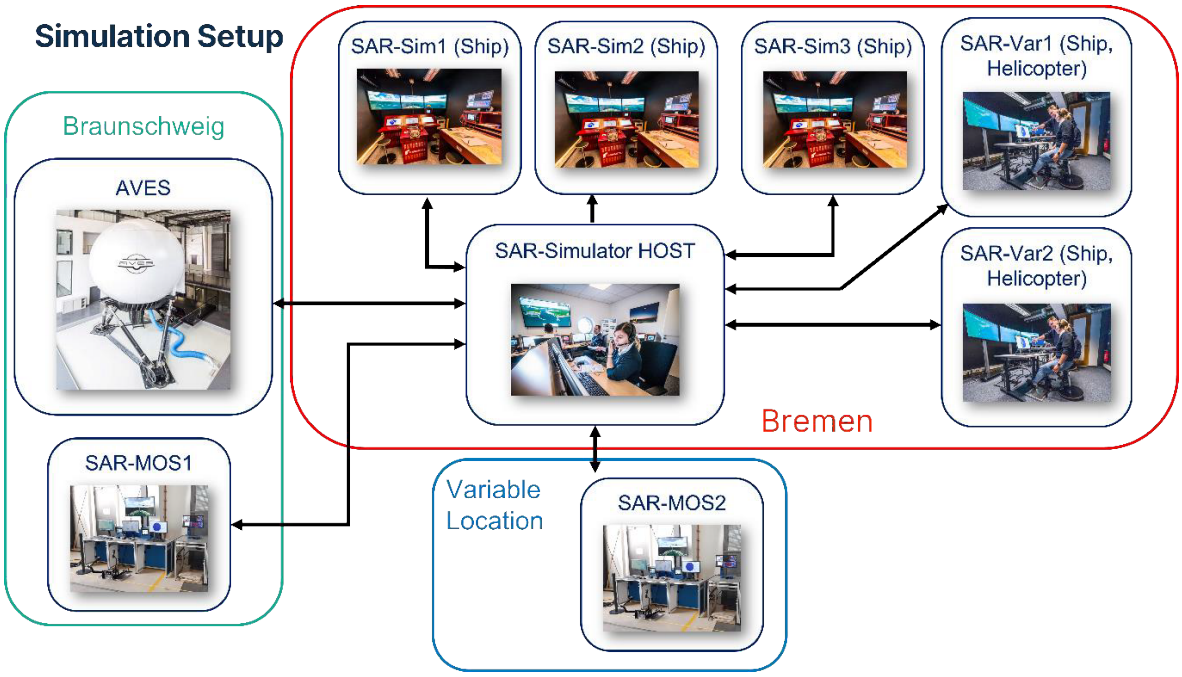


Figure 1. Simulation Setup

With the realization of this innovation, ship crews can train together with helicopter crews in joint scenarios—they can see the other units and have all the means of communication available to them on their real rescue units. In our quantitative study, we evaluate the effectiveness of our training in this innovative simulation environment, by collecting survey data on participants' perceived training success using the Feedback Instrument for Rescue Forces Education (Thielsch & Hadzihalilovic, 2020). This data shows that our training has a positive effect on the acquisition and transfer of competencies as rated by the participants. Additional analysis shows that the extent to which individuals benefit from the training varies based on professional background, age, and tenure. This provides criteria for selecting future training participants to maximize the potential future benefits of the training.

OPERATIONAL BACKGROUND

The close practical relevance of our research has its origins in the operational practice of the maritime search and rescue service in Germany. The authority for this was transferred from the Federal Ministry of Transport to the non-profit organization DGzRS. This organization not only coordinates the rescue operations but also provides the necessary ships (IMO, 2006). In Germany, air support at sea is regulated by a mutual support agreement with the Ministry of Defense, which is responsible for the aeronautical search and rescue service. This means that SAR helicopters from the German Navy are also available on request of the DGzRS for rescuing people in distress at sea. Furthermore—and because international conventions oblige anyone in the vicinity of a maritime emergency to provide assistance—private offshore aviation represent a potential additional resource for DGzRS maritime rescue operations. With the ongoing expansion of offshore wind farms in the German Bight and the western Baltic Sea, these companies are now also present in the German Exclusive Economic Zone. The helicopters of these companies transport people and material to the wind farms or back to shore, and usually have medical personnel on board. The increasing economic use of maritime space for the production of renewable energies therefore generates an increase in the number of ships and helicopters that could assist in rescue operations. Formally, the tactical interaction between these resources in a rescue situation is regulated by an international set of rules. Specifically, International Aeronautical and Maritime Search and Rescue Manual (IAMSAR) published by the International Maritime Organization (IMO) and the International Civil Aviation Organization (ICAO) describes all essential procedures for the search and rescue of human life and is the common action framework for SAR organizations in an international context (IMO & ICAO, 2022). The organizations with a SAR mission—in Germany, the DGzRS and the Naval Air Wing 5—use the IAMSAR Manual as a legally binding tactical interaction framework. In addition, the officers of every merchant ship are required to have a current edition of

the IAMSAR Manual at hand, in case their ship finds itself in distress at sea or has to carry out search and rescue operations on site. The situation is somewhat different in aviation. In civil aviation, the IAMSAR Manual does not have to be carried on board, and it is largely unknown to pilots. In the past, this has been repeatedly observed in the operational integration of offshore helicopters in rescue missions and real live exercises: Although offshore helicopters demonstrate continuously excellent skill at rescuing people in distress from ships or from the water, we found deficits in carrying out search operations because the necessary know-how is not mandatory for helicopter pilots. Taken together, these circumstances provide an excellent setting to advance our understanding of training for inter-professional ad hoc teamwork.

Our training is implemented by the DGzRS simulator training crew. For more than 20 years, the DGzRS simulator crew has been training nautical officers and captains of government vessels, in addition to its own emergency crews, to be able to take on leadership roles during maritime SAR operations (e.g., on-scene coordinators [OSC]). Our study draws on this extensive experience. Specifically, the AMARIS multi-day training course mainly takes place at the DGzRS SAR simulator center with participants from different organizations/authorities, with the HEMS professionals joining the training from the DLR facility in Braunschweig. To operationalize ad hoc teams, we randomly allocate seafaring professionals to trainings teams of three. These participants each occupy a simulator cubicle and apply the course content in scenarios of varying complexity. The HEMS professionals in Braunschweig were not subject to this randomization in team allocation, because there is only one cubicle in Braunschweig.

METHODS

In order to develop and test our training concept for inter-professional (seafarer and HEMS professionals) ad hoc teams, we conducted a two-phase study. First, to identify requirements both for such a training concept and a coupled simulator environment, we conducted an exploratory qualitative data collection using 25 semi-structured, guided interviews. Second, to test the effectiveness of our training concept, we enrolled a preliminary number of 49 unique participants (40 seafaring professionals, 9 HEMS professionals) and subjected them to the training protocol using the linked DGzRS and DLR simulators. In this phase of our data collection, we asked the participants to complete a survey after completing the training. The following paragraphs provide a detailed account of both phases of data collection and analysis.

In study phase one—our interview study—we approached individuals knowledgeable about the phenomenon we intended to study: We interviewed SAR stakeholders, such as helicopter pilots (13) and coxswains of rescue cruisers (8), but also safety officers (2), trainers (1), and training representatives from the participating organizations (1). We decided to use the data collection method of semi-structured interviews, because they offer the advantage of providing comparable responses across cases but can also be adapted to the background and experiences of the interviewees, thus allowing for additional exploration if needed. This is because semi-structured interviews allow for follow-up questions on individual aspects and any ambiguities can be clarified immediately (King, Horrocks, & Brooks, 2019). The interviews were conducted virtually[3] and were recorded and transcribed in accordance with both methodological (Miles, Huberman, & Saldaña, 2020) and data protection guidelines. The interview guideline was developed interdisciplinary, taking into account aeronautical, maritime SAR, and behavioral science perspectives at the same time. It covered items such as description of a recent and demanding incident concerning procedure, key aspects of ship – helicopter teamwork in SAR operations, and assessment of occurring challenges and how they were solved. To provide specific instances where inter-professional ad hoc teamwork is needed, special emphasis was placed on the consideration and description of various hoisting procedures and typical challenges and possible misunderstandings between helicopter pilots and coxswains when carrying out a hoisting maneuver. Another key focus was put on the technical set up of a joint simulation environment. Helicopter pilots as well as coxswains were asked to describe key features which they would like to see in a shared simulation environment. This information was essential to our technical development of a linked simulator environment. Overall, we collected 33 hours of interview data. On average, each interview lasted 78 minutes (min: 48; max: 113). We processed our interview data using content analysis. Our analysis of the qualitative data revealed that although the IAMSAR Manual Vol. III (IMO & ICAO, 2022)—as international SAR standard—regulates the cooperation between seafaring and flying units during missions, relevant regulations were not always sufficiently well known by interviewees, especially among commercial offshore pilots. Flying and seafaring units partly show significant differences in standard operation procedures and in communication routines. We found that this can have a negative impact on the quality of cooperation when carrying out both standard search procedures and hoisting maneuvers. For example, interviewees reported that when conducting

[3] We conducted the interviews virtually due to the COVID pandemic. To ensure interoperability and good data quality, we used the Microsoft Teams platform and its recording capabilities.

hoisting maneuvers, they found it essential that the processes run smoothly and that everyone involved can rely on each other. We also found that although these operations were often practiced under real conditions, the interviews expressed low confidence in their knowledge of the working conditions of the other unit (helicopter or rescue cruiser). Thus, the main result of study phase one was the identification of specific challenges and needs for improvement of joint SAR operations involving helicopters and rescue cruisers. The interview data also provided requirements for realizing a linked simulation environment for aeronautical and maritime SAR training, which informed our technical implementation of the simulator coupling.

We used the findings from study phase one, as well as additional expert workshops, to derive the training concept ("AMARIS Training"). Two workshops involving various SAR experts with either maritime or aeronautical background were conducted to identify and refine both training objectives and corresponding learning objectives. The first expert workshop started with defining the *key objective* of the AMARIS training ("Coordination of multiple [air and sea] units to manage a maritime SAR situation on the basis of IAMSAR III"). On this basis, the learning objectives and the training content to be derived from them were developed gradually and with an increasing level of detail. Using the example of the training objective "Raising awareness of the capabilities and characteristics of seagoing and flying units", the training content includes, for example, "imparting knowledge about the definition of operational limits", "recognizing dangerous situations in practice", and "reacting appropriately to the use of an emergency procedure by the other unit". All learning objectives and training content have to be considered both from an intra-profession and a cross-profession perspective. This means though the AMARIS training aims to improve ship-helicopter interaction, training effects in intra professional cooperation can also be expected.

The second workshop was centered on developing the AMARIS simulator scenarios. These were designed with a focus on the AMARIS learning objectives as well as the knowledge and skills previously identified. We collaborated with groups of aeronautical and maritime SAR experts, with the assistance support of the SAR simulator trainer team, to develop three scenarios. These scenarios varied in task complexity, starting from a less-complex "starter scenario" to a highly-complex major incident. We conducted multiple check-ups to ensure that all previously identified training objectives, knowledge, and skills were incorporated and trained in these scenarios. Once the training scenarios were finalized, we created written training materials. These training materials covered basic and advanced SAR knowledge (e.g., search patterns) as well as information on the special objectives of the AMARIS training.

Finally, an extensive testing and programming work took place in parallel with the content design of the training program to ensure the technical framework conditions (Szczygielski, Gerten, & Wilhelm, 2021).

Procedures and Participants

In study phase two—the testing of the effectiveness of the AMARIS training—we implement the AMARIS training using seafaring and HEMS professionals. We first provide information on the course of the training before describing our study sample. The AMARIS training contains two elements: First, an initial online information session with participants and representatives of the AMARIS project. This session was organized to ensure that both groups of participants (seafaring and HEMS professionals) were well-informed about the theoretical SAR basics, the specific objectives of the AMARIS training, and the research data collection. The session also provided details on data protection and allowed future participants to clarify related questions to ensure informed consent. Second, the AMARIS training in the two linked simulators (DGzRS in Bremen, DLR AVES in Braunschweig). Each training starts with signing the data protection agreement, followed by blocks of theoretical and practical instructions in data collection and simulator handling. Hybrid sections via a virtual communication platform alternate with separate instructions in Bremen and Braunschweig. All scenarios start with a joint virtual briefing and conclude with a virtual debriefing session connecting all teams in Bremen and Braunschweig. In the default mode, each training sessions covers three ship cubicles with teams of three seafaring professionals each in the SAR simulator in Bremen and three HEMS professionals (pilot, co-pilot, hoist operator) in the AVES simulator in Braunschweig, thus matching our theoretical interest inter-professional work. To ensure that (most of) our study participants operated in ad hoc teams, participants in the ship cubicles switched crew compositions for every scenario. As there is only one helicopter crew in the AVES simulator, this crew composition remained constant. Please note however that the inter-professional collaboration between ships—due to their changing crew composition—and the helicopter also operationalizes ad hoc teamwork. The simulator was operated by a team of 3-4 instructors with the support of 1-2 representatives of the Aeronautical Rescue Coordination Centre Glücksburg, Germany (ARCC). These individuals oversaw the training from the instructor room in Bremen, simulating foreign ships, the MRCC, and the ARCC. Data collection, storage, and analysis were performed exclusively by the university research team in accordance with the study's data protection protocol. We pretested the AMARIS training and data collection three times using instructors from sailing and water sports

schools. These pretests resulted in minor modifications of the study protocol.

At the time of writing, 50 participants from various organizations participated in five AMARIS training sessions. Because one participant had to leave the AMARIS training early due to professional commitments, our preliminary sample consists of 49 individuals. Seafaring professionals came from the DGzRS (permanent [17] and volunteer [7] sea rescuers), further participants were also professional seafarers with theoretical and practical expertise in participating in sea rescue operations (16). Helicopter pilots came from the Navy (6), from (commercial) offshore helicopter service companies (2), and from general air rescue service providers (2). The heterogeneity in participants' organizational backgrounds provides an ideal match with our interest to study inter-professional ad hoc teams. On average, participants had worked for their organizations for 9 years (SD = 9.26) and in their jobs for 6 years (SD = 8.07). The mean age was 44 years (SD = 11.22). The gender distribution of 45 (91.84%) males is consistent with the gender distribution in the seafaring profession.[4]

Measures

We collected data on study participants' demographics, professional background, and—after completing the AMARIS training—perceived training effectiveness. These data were collected using a survey. Prior to the data collection, we discussed the survey with representatives of the professional groups involved in the AMARIS study, and we pretested scales with individuals who did not participate in the main study. This pretest consisted of a think aloud protocol to receive structured feedback on the validity and comprehensibility of the scales (Sudman, Bradburn, & Schwarz, 2010).

Outcome. To measure the perceived effectiveness of the AMARIS training, we asked participants to complete the Feedback Instrument for Rescue Forces Education, which we provided as a validated German translation (Thielsch & Hadzihalilovic, 2020). This instrument allows participants to subjectively assess their perceived training success, providing important information on the effectiveness of the AMARIS training. To ensure valid responses, the wording of the items was slightly adapted from the instrument's original firefighting context to our SAR context. Overall, the instrument covers six subscales—Trainer Behavior, Overextension, Structure, Group Behavior, Competence Acquisition, and Transfer—relevant to the AMARIS training evaluation. Answers ranged from 1 ("strongly disagree") to 7 ("strongly agree"). For each subscale, the corresponding items formed a single scale (Trainer Behavior: Cronbach's $\alpha = 0.71$, Overextension: Cronbach's $\alpha = 0.92$, Structure: Cronbach's $\alpha = 0.90$, Group Behavior: Cronbach's $\alpha = 0.83$, Competence Acquisition: Cronbach's $\alpha = 0.96$, Transfer: Cronbach's $\alpha = 0.75$). For this preliminary report, we focus on the two outcome subscales, i.e., self-rated Competence Acquisition and Transfer (Blanchard & Thacker, 2023).

Independent Variables. To identify which individuals perceived to benefit more—and which ones less—from the AMARIS training, we focused on established demographic and job predictors of training outcomes (Nazir, Jungefeldt, & Sharma, 2019; Sellberg, 2017; Colquitt, LePine, & Noe, 2000). Specifically, we tested for differences in the outcome for the following grouping variables: Age, Tenure (Organization), Tenure (Job). To create grouping variables for these variables, we conducted median splits. This procedure resulted in the following independent grouping variables: Age (younger: 25 years – 42 years, older: 43 – 69 years), Tenure Organization (low tenure: 0 – 5 years, high tenure: 6 – 32 years), Tenure Job (low tenure: 0 – 3 years, high tenure: 4 -32 years), and trainees' formal Education (1 = Secondary School, 2 = Intermediate Maturity, 3 = University Entrance Qualification, 4 = Graduate). We do not examine trainee gender due to the limited variance in this demographic in our current sample. However, we tested whether DGzRS Status (0 = non-DGzRS, 1 = DGzRS Volunteer, 2 = DGzRS Employee) explains differences in the outcome variables.

RESULTS

All analyses presented below were performed in Stata 16.1. Table 1 displays descriptive statistics and correlations among our study variables. The directions of the significant correlations are in the expected directions. In addition, as expected, we observe strong correlations between tenure organization and tenure job ($r = 0.72$). Since we do not present a multivariate analysis in this preliminary report, this unsurprising correlation is not a cause for concern.

[4] According to the Baltic and International Maritime Council & the International Chamber of Shipping (2021), women worker represents 1.28% of the global workforce in the shipping industry. The International Transport Workers' Federation (2024) states, the global proportion of female workers in the maritime sector is 2%.

Table 1. Means, Standard Deviations, and Correlations

	Mean	SD	1	2	3	4	5	6
1. Age*	1.47	0.50						
2. Tenure Organization*	1.49	0.51	0.39					
3. Tenure Job*	1.43	0.50	0.34	0.72				
4. Education	3.35	0.80	-0.26	0.03	0.19			
5. DGzRS Status	0.84	0.92	0.35	0.40	0.34	-0.29		
6. Competence Acquisition	4.38	1.58	0.29	-0.31	-0.05	-0.10	0.12	
7. Competence Transfer	4.95	1.32	0.16	0.15	0.41	-0.01	0.28	0.38

n = 49, correlations larger than |0.29| are significant at $p < 0.05$.
* Dummy variables based on median-split (1 = low, 2 = high).

To show the extent to which trainees perceived the AMARIS training to benefit their competence acquisition and transfer, Table 1 provides corresponding descriptive statistics. Table 1 shows an average Competence Acquisition of 4.38 (out of 7). The standard deviation for Competence Acquisition is 1.58, suggesting relevant variance in the participants' perception. The participant's average Competence Transfer (4.95) exceeds average Competence Acquisition, and a slightly smaller standard deviation (1.32), suggesting that the training offers relevant insights for advanced professionals with practical expertise. To better understand which trainees perceive that they benefit more (or less) from the AMARIS training, we performed the non-parametric Wilcoxon rank-sum tests for all binary independent variables. For the two categorical independent variables (Education, DGzRS Status), we conducted Kruskal-Wallis tests.

For Competence Acquisition, we found that trainees with higher Age had a significantly higher level than trainees with lower Age (p = 0.05). Figure 2 provides a strip plot illustrating this result. In contrast, trainees with low Tenure in their Organization had a marginally significant higher level of perceived Competence Acquisition than trainees with high Tenure in their Organization (p = 0.06). Figure 3 provides a strip plot illustrating this result. Surprisingly, we observe that trainee Tenure in the Job does not appear to explain any of the variance in this outcome (p = 0.83). We also do not observe significant differences in perceived Competence Acquisition for different levels of Education (p = 0.81) or DGzRS Status (p = 0.82).

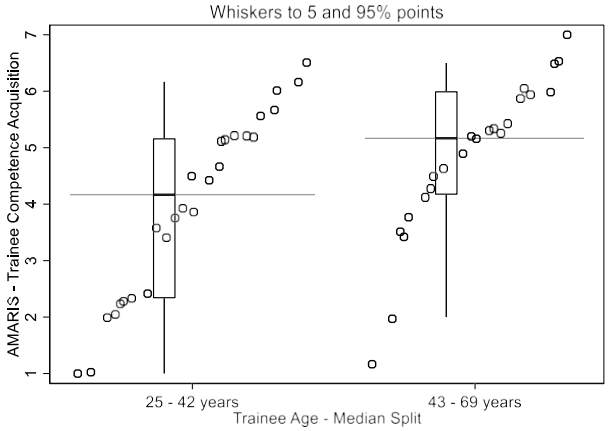


Figure 2. Competence Acquisition and Age



Figure 3. Competence Acquisition and Tenure Organization

For Competence Transfer, we found no significant difference between trainees with higher and lower Age (p = 0.32) and Tenure in their Organization (p = 0.21). However, our results show that trainees with high Job Tenure had significantly higher levels of perceived Competence Transfer than trainees with low Job Tenure (p = 0.01). Figure 4 provides a strip plot illustrating this result. We do not observe significant differences in perceived Competence Transfer for different levels of Education (p = 0.45) or DGzRS Status (p = 0.11).



Figure 4. Competence Transfer and Tenure Job

Taken together, the results for Competence Acquisition are consistent with the notion that older trainees with limited exposure to their current organization perceive themselves as particularly benefiting from the AMARIS training. The results for Competence Transfer are consistent with the argument that trainees with more extensive work experience in the maritime or aviation environment have more extensive mental representations of situations to which they can relate the AMARIS training material and experiences. In addition, as suggested by the non-significant results for both outcomes and Education and DGzRS Status, the AMARIS training does not appear to be exclusive in terms of formal education or professional SAR affiliation, suggesting that the training is inclusive in nature.

DISCUSSION AND CONCLUSION

The data available to us so far shows that the AMARIS training tested in our study is perceived as effective as a cross-professional qualification program. This is a promising initial result, suggesting that the previously identified training needs and critical mission situations cover relevant issues, and that the AMARIS training effectively addresses those issues. We were able to demonstrate that participants perceived that they gained skills

from the simulator training, regardless of their age and professional experience. However, the results need to be scrutinized closely, as we found that older participants, in particular, gained more competencies overall than younger training participants.

This is significant as older professionals are often difficult to reach with simulator training courses (Gully & Chen, 2010) as these professionals often do not consider such courses to be ”real live” or “hands-on“. Two conclusions can be drawn from this, although additional data and analyses are required: First, our simulators may have created a level of immersion that allowed all participants to perceive the situations created in the scenarios as authentic and to engage with them. This would be positive a training-inherent effect, as it could provide effective starting points for inclusive simulator training.

Second, the older participants may have a deeper understanding of the factual necessity of specifically training the topic of cross-professional collaboration (e.g., “effective ship-helicopter interaction”), so that potential concerns or reservations about simulator training may have simply faded. We therefore find a greater awareness of the importance of knowledge among the older study participants, which indicates an increasing level of professionalization over the course of their professional lives.

A surprising result for us was the significant correlation between the length of time spent in the organization—i.e., Tenure: Organization—and Competence Acquisition: participants who have been with their current organization for longer showed a lower level of competence acquisition than people who have only recently started working for their organization. The question of whether this could be due to motivational- or competency-related causes affected by the organizational tenure would have to be investigated using a larger sample, multivariate methods, and potentially additional data sources.

PRACTICAL IMPLICATIONS

The cross-domain networking of two different simulators implemented in the AMARIS project offers the possibility of linking the SAR Simulator Center with other flight simulators. This will allow the implementation of joint training content and new transdisciplinary training formats in the future, but it will also open up new research and development opportunities: The cross-disciplinary simulation of complex rescue missions suggests to be a promising laboratory environment for testing communication paths, operational procedures, and information systems with emergency personnel and obtaining well-founded end-user feedback. the possibility of operating prototypes of newly developed assistance systems during the simulation show great potential for researchers to bring innovation and end users together at an early stage.

LIMITATIONS

Our study does have limitations. Some will be addressed in our full manuscript, while others may inspire future research. First, this preliminary study analyzes a relatively small data set. To provide a robust analysis with our current number of observations, we rely on simple statistical tests. We are aware that we are currently using binary predictors (by means of median splits), and we cannot use control variables to account for potential omitted variable bias. While seafarers were randomly allocated to teams, our study is not an experiment, as it does not have control and experimental conditions. Thus, using multivariate methods once our data set is sufficiently large is our next step to rule out such issues. Second, due to limitations in both the simulator setup and the availability of helicopter pilots, we could not randomize team allocations among helicopter staff within each experiment. Future research should attempt to overcome this shortcoming by operating in settings with sufficiently large populations to ensure full randomization of participants.

ACKNOWLEDGMENTS

The joint research and development project AMARIS is part of the Security Research Program of the German Government and funded by the German Federal Ministry of Education and Research (BMBF), grant no. 13N15500, 13N15501, and 13N15502.

REFERENCES

Baltic and International Maritime Council & International Chamber of Shipping (2021). *Seafarer workforce report 2021*. July 2021.

Birkmann, J., E. Liwenga, R. Pandey, E. Boyd, R. Djalante, F. Gemenne, W. Leal Filho, P.F. Pinho, L. Stringer, & Wrathall, D. (2022): *Poverty, livelihoods and sustainable development*. In: *Climate Change 2022: Impacts, adaptation and vulnerability. Contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change* [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 1171–1274. https://doi.org/10.1017/9781009325844.010

Blanchard, P.N., & J.W. Thacker (2023). *Effective training: Systems, strategies, and practices*. (7th ed.) Thousand Oaks: SAGE.

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *The Journal of Applied Psychology, 85*(5), 678–707. https://doi.org/10.1037//0021-9010.85.5.678

DLR – German Aerospace Center (2024). *AVES motion simulator.* https://www.dlr.de/de/forschung-und-transfer/forschungsinfrastruktur/grossforschungsanlagen/aves-motion-simulator, (2024-04-12).

Flin, R. H., O'Connor, P., & Crichton, M. (2007). *Safety at the sharp end: Training non-technical skills.* Aldershot: Ashgate.

Gully, S & Chen, G. (2010). Individual differences, attribute-treatment interactions, and training outcomes. In: Kozlowski, S. W. J., & Salas, E. (2010). *Learning, training, and development in organizations. The organizational frontiers series*. New York: Routledge, pp. 3–64.

Hällgren, M., Rouleau, L., & de Rond, M. (2018). A matter of life or death: How extreme context research matters for management and organization studies. *Academy of Management Annals, 12*(1), 111–153. https://doi.org/10.5465/annals.2016.0017

Hollenbeck, J.R., Beersma, B., & Schouten, M.E. (2012). Beyond team types and taxonomies: A dimensional scaling conceptualization for team description. *Academy of Management Review* 37(1), 82-106.

International Maritime Organization & International Civil Aviation Organization (2022). *International aeronautical and maritime search and rescue manual: IAMSAR Manual Vol. III: Mobile facilities*. *IMO publication*. London: IMO.

International Maritime Organization (2015). *GMDSS Manual: Global maritime distress and safety system* (8th ed.). *IMO publication*. London: IMO.

International Maritime Organization (2014). *SAR on-scene coordinator (IAMSAR Manual, Volume III)* (2014 ed.). *Model course: Vol. 3.15*. London: IMO.

International Maritime Organization (2006). *SAR convention: International convention on maritime search and rescue, 1979; as amended by resolutions MSC.70(69) and MSC.155(78); 2006 edition* (3rd ed.). *IMO publication*. London: IMO.

International Transport Workers' Federation (2024). *Women seafarers*. https://www.itfseafarers.org/en/issues/women-seafarers, (2024-02-07).

King, N., Horrocks, C., & Brooks, J. M. (2019). *Interviews in qualitative research* (2nd ed.). Los Angeles: SAGE.

Klein, G. (2015). A naturalistic decision making perspective on studying intuitive decision making. *Journal of Applied Research in Memory and Cognition, 4*(3), 164–168. https://doi.org/10.1016/j.jarmac.2015.07.001

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. Upper Saddle River, NJ: Pearson Education.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). Los Angeles: SAGE.

Nazir, S., Jungefeldt, S., & Sharma, A. (2019). Maritime simulator training across Europe: A comparative study. *WMU Journal of Maritime Affairs, 18*(1), 197–224. https://doi.org/10.1007/s13437-018-0157-0

Salas, E., DiazGranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008). Does team training improve team performance? A meta-analysis. *Human factors, 50(6),* 903-933. https://doi.org/10.1518/001872008X375009

Shippmann, J. S., Ash, R. A., Batjtsta, M., Carr, L., Eyde, L. D., Hesketh, B., Kehoe, J., Pearlman, K., Prien, E. P., Sanchez, J. I. (2000). The practice of competency modeling. *Personnel Psychology, 53(3)*, 703-740.

STM – Sea Traffic Management (2024). https://www.seatrafficmanagement.info, (2024-04-12).

Sellberg, C. (2017). Simulators in bridge operations training and assessment: A systematic review and qualitative synthesis. *WMU Journal of Maritime Affairs, 16*(2), 247–263. https://doi.org/10.1007/s13437-016-0114-8

Steigenberger, N. (2016). Organizing for the big one: A review of case studies and a research agenda for multi-agency disaster response. Journal of Contingencies and Crisis Management, 24(2), 60–72. https://doi.org/10.1111/1468-5973.12106

Sudman, S., Bradburn, N.M., & Schwarz, N. (2010). *Thinking about answers*. San Francisco: John Wiley.

Szczygielski, T., Gerten, E., & Wilhelm, H. (2021*). AMARIS Schulungskonzept.* Universität Witten/Herdecke.

Thielsch, M. T., & Hadzihalilovic, D. (2020). Evaluation of fire service command unit trainings. *International Journal of Disaster Risk Science, 11*(3), 300–315. https://doi.org/10.1007/s13753-020-00279-6

Weick, K. E. & Sutcliffe, K.M. (2015). *Managing the unexpected: Sustained performance in a complex world.* Hoboken: Wiley.

Weick, K. E., Sutcliffe, K.M., & Obstfeld, D. (1999). Organizing for high reliability: Processes of collective mindfulness. *Research in Organizational Behaviour, 21*, 81-123.

Weick, K. E., & Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly, 38*(3), 357-381. https://doi.org/10.2307/2393372

User-tailored visualization of simulation and sensor data for efficient crisis management

Nils Winter[1,4]
nils.winter@iosb.fraunhofer.de

Michael Monteforte[1,3]
michael.monteforte@ivi.fraunhofer.de

Fabian Müller[1,2]
fabian.mueller@emi.fraunhofer.de

Ines Rohrbach[1,3]
ines.rohrbach@ivi.fraunhofer.de

Philipp Hertweck[1,4]
philipp.hertweck@iosb.fraunhofer.de

Till Martini[1,2]
till.martini@emi.fraunhofer.de

[1]Fraunhofer Center for the Security of Socio-Technical Systems SIRIOS, Berlin, Germany

[2]Fraunhofer Institute for High-Speed Dynamics, Ernst-Mach-Institute, EMI, Freiburg, Germany

[3]Fraunhofer Institute for Transportation and Infrastructure Systems IVI, Dresden, Germany

[4]Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Karlsruhe, Germany

ABSTRACT

The availability of all relevant data in a situation report is key for efficient crisis management. The digital twin concept allows for comprehensive data accumulation, both from real-world sensors as well as augmented by simulations. However, the excessive amount and detail of data can hinder usability and discourage users or impede efficient decision making. To still benefit from such emerging technologies, user-tailored data visualization with appropriate filtering and processing is recommended. This work promotes the use of a standardized data aggregation in order to optimize customized data processing through implementation into existing platforms for crisis management. The concept is showcased for three well-established platforms with data accumulated from a standardized digital twin representation of urban critical infrastructure and emergency units. Each of the examples focuses on different facets of visualization as typically demanded by the respective target user group.

Keywords

Data aggregation, user-tailored data visualization, simulation and sensor data, crisis management

INTRODUCTION

The increasing frequency and intensity of critical situations resulting from both man-made and natural disasters require a heightened focus on security and resilience measures. A major focus of these measures is on critical infrastructure, as the failure or disruption of this infrastructure can have a significant impact, potentially leading to supply shortages and failures, thus reducing the level of public safety. Due to the highly crosslinked nature of the modern world, the shortfall of one single infrastructure provider not only has an impact on the immediate

environment but might also affect other critical sectors. In particular, unanticipated interdependencies and cascading effects between different critical infrastructures may lead to the increased risk of severe disruptions (Rehak et al., 2016; Rehak et al., 2018). Therefore, a deep understanding of interdependencies and cascading effects is required in order to take targeted actions and maintain public safety.

The high level of interconnectedness of critical infrastructure also has an impact on the assessment of a critical situation: Knowledge of a single domain is often not sufficient to respond to the situation, as in many cases several sectors are involved. Nevertheless, it is crucial for decision makers to act quickly and efficiently to prevent worst case scenarios. In such situations, visualization is a key tool to break down the complexity and simplify data access and interpretation. It helps stakeholders to understand, analyse, and respond effectively to complex scenarios, leading to better outcomes and risk mitigation. Furthermore, it enables people from outside the field to gain understanding and interpretation of the situation.

Large sensor networks and shared data bases, especially when augmented by computer-aided modelling and numerical simulations, can provide useful but potentially overwhelming amounts of information about complex situations. For analyses regarding the aforementioned interdependencies and cascading effects, disentanglement and preprocessing of information is needed.

The well-known proverb “A picture is worth a thousand words” underlines the power of a visualization. However, the quality of the visualization is also crucial, as every simplified visualization always runs the risk of leading to misunderstandings and misinterpretations. In order to enhance the correct understanding and assessment of a critical situation, it is necessary to adapt the visualization to the needs of different user groups.

This paper presents the concept of a versatile processing pipeline for large comprehensive amounts of complex, interdependent data using three different visualization tools. All of them are able to highlight relevant aspects of critical situations to their respective user groups. This is made possible by the consistent use of open standards as an intermediary between data producers and consumers. The three compared tools are a representative selection of well-established tools used by different user groups in public safety. However, other visualization software could be used as well.

The chosen tools were all developed at Fraunhofer which provided easy access to necessary implementations by the authors. The *Elektronische Lagedarstellung* (ELD, engl. *Electronic Situation Dashboard*) is used extensively in the context of operations and exercises at local governments in Germany. The *Digitaler Lagetisch* (DigLT, engl. *Digital Map Table*) focuses on shared situation visualization and analysis in civilian and defense applications. The targeted user groups of the operational command software *MobiKat*® range from the command staff to the relief forces of fire and rescue services, as well as disaster managers and the police. All these user groups work on public safety from different viewpoints and have different requirements which led to these different software systems. This paper presents an approach to visualize complex data in these established software systems tailored to each user group.

BACKGROUND

The Fraunhofer Center for the Security of Socio-Technical Systems (SIRIOS) builds a research, testing and training environment based on coupled simulations of complex security scenarios. The developed co-simulation framework is presented in Martini et al. (2024). It is designed to explore and assess the impact on citizens and to be used for training purposes or in real-life situations to evaluate different counter measures to security threats. The framework provides a digital twin representation of relevant critical infrastructure sectors in an urban environment. Special attention is paid to reproducing and disclosing dynamic cascading effects between the respective sectors through a coupled simulation with defined mutual interdependencies. To this end, stand-alone modules for the simulation of several critical infrastructures are included, e.g., power, gas, water, and telecommunication grids as well as the operational dynamics of emergency response units and hazard-dependent vulnerability assessment of building structures. The framework is an ideal information source for the intended processing and visualization pipeline as it offers the ability of incorporating a large variety of heterogeneous simulation modules which are based on, e.g., engineering models rooted in physical principles, on agent- and rule-based modelling or on structural engineering analyses. Especially in conjunction with available real-life sensor data, this approach maximizes the information content deliverable by the digital twin. Thus, it can provide large amounts of interdependent data.

Naturally, each of the incorporated modules produces individual formats of output data. Utilizing well-established concepts from the *IoT* (internet of things) sector together with standardized data models proves to have versatile benefits. The Open Geospatial Consortium (OGC), a neutral and collaborative platform within the geospatial industry fostering innovation and development of open standards (van Rees, 2013), has introduced such a model:

the *OGC SensorThings API* (Liang et al., 2021). It offers a canonical way for embedding arbitrary heterogeneous modules into the framework if their entities and results can be interpreted within the respective *IoT* data model. To reach a level of data harmonization, the presented framework suggests interpreting the key entities of each infrastructure model as *Things*[1] equipped with *Sensors* which provide synthetic *Observations* from the respective simulation results to be stored on a *FROST* server instance (Hertweck et al., 2019; Fraunhofer IOSB FROST, 2023). Additionally, the interdependencies between different infrastructures can be defined on the basis of the thing representation of their respective models. Thereby, it is specified how the inter-modular information transfer required for the coupled simulation can be realized independently of the inner details of the actual simulation modules. Treating simulated and actual sensor data on an equal footing maintains the concept of the digital twin when incorporating simulation results in the accumulated data.

Each of the coupled simulation modules potentially produces an arbitrary amount of data, which might not be completely relevant in all details for the specific user demand of the digital twin framework. Thus, preselection and data aggregation are necessary for optimal, efficient use of the provided data. Assuming applications of the framework in training settings of relief units and task forces, or infrastructure managers or as aid for decision support used by decision makers, authorities, and organizations with security responsibilities, implies human users with diverse professional backgrounds. Besides appropriate preselection and aggregation, suitable and flexible visualization options of the data are therefore mandatory in order to concisely convey relevant information as efficiently as possible.

An increasing number of municipalities, service providers as well as emergency units rely on well-established digital tools like dashboards, geographic information systems (GIS) and even digital twin models to aggregate, display and analyze data. In order to seamlessly integrate the results of the co-simulation framework into such already existing eco systems and to benefit from well-rehearsed workflows for analysis and visualization of the data, providing open standard interfaces is advantageous. This is especially important for retaining an accustomed level of usability together with a low entry threshold after the introduction of new technology into a fast-paced, heterogeneous environment like crisis management. Data representation and storage as well as connecting data entities in the digital twin representation provided by the co-simulation framework relies on the *OGC SensorThings API* standard. Thus, further data processing and analysis followed by suitable visualization can, in principle, be carried out by any tool which is compatible with this *IoT* standard (Russell et al, 2018) (see figure 1). Utilizing the comprehensive digital twin representation as the central information backbone with a standardized interface to various visualization and analysis platforms facilitates additional benefits regarding resource management: Laborious tasks like complex simulation runs as well as data acquisition and aggregation need only to be performed once in a suitable, possibly remote, environment like, e.g., a high-performance computing cluster. Different user groups can subsequently access and update, if possible, specifically funneled data quickly and in parallel over their respective platforms of choice (see figure 1). Furthermore, this pipeline ensures a common and synchronous database through all involved platforms and tools. Note that an update of the common database by one tool might trigger a renewing simulation run with implications for all online tools (as indicated by the reverse arrows in figure 1).

VISUALIZATION TOOLS

In the following, the three representative visualization and processing tools, which are commonly used by the foresaid user groups, are briefly showcased. Their individual properties and features are highlighted together with accordingly customized filtering, summarizing, processing, and visualization options of the data to accomplish efficient data transfer tailored to the respective target user group.

[1] In the *OGC SensorThings API standard*, a *Thing* is an arbitrary *IoT* device or system able to provide streams of sensor data.

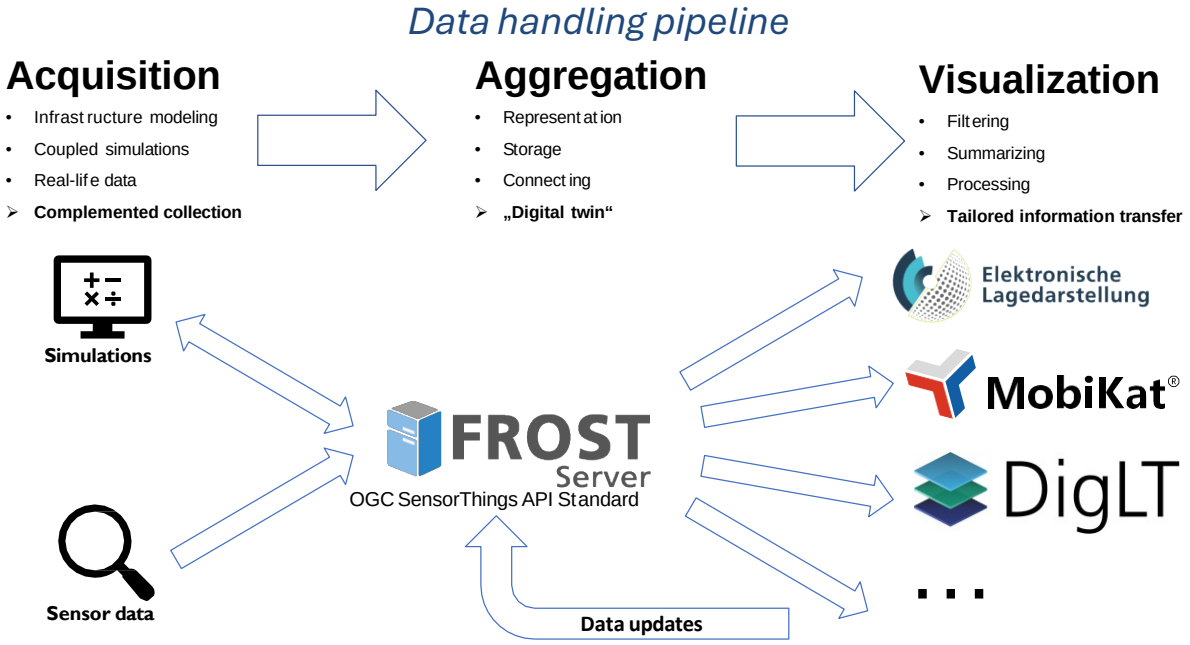


Figure 1. From data acquisition to user-tailored visualization.

ELD

The *Elektronische Lagedarstellung* (ELD, engl. *Electronic Situation Dashboard*) is a web-based information system that links a wide range of data to support decision making. One of its main purposes is to support staff work and crisis management by authorities in critical situations. The software enables the collection, management, and visualization of raw data (e.g., time series from sensors) as well as expert knowledge. The ELD is based on open standards. This allows for easy incorporation of external data sources. It demonstrates its strengths when the given data has a spatial reference as the integration of maps is one key function. Supported formats include Web Map Service (WMS) and *IoT* sensor data such as *OGC SensorThings API* (Fraunhofer IOSB, 2024).

As described in Martini et al. (2024), the proposed co-simulation framework relies on the concept of digital twins – and simulation data can be seen as sensor data. Therefore, the ELD interface to the *OGC SensorThings API* with special support for the open-source product *FROST* allows an integration and visualization of the simulation data. In general, all *Things* on a *FROST* server can be displayed in the ELD, regardless on their level of granularity. Pre-processing of all available data is therefore necessary in order to reduce the amount of data and generate a user-tailored visualization.

The ELD map relies on the concept of individual visual layers. This makes customization of the visible data easy: Layers can be overlaid to aggregate information and hidden to filter information. To this end, the complete result of the co-simulation is broken down to the results of single components. These individual results can in turn be displayed on one or more map layers. Each map layer can then be individualised by defining custom icons, different colours and further settings, e.g., line style. In addition, the use of layer groups makes it possible to preset the visualization for different user groups.

One group could include all data on the public electricity grid for an infrastructure operator, from substations to individual power lines and consumers. Another could contain relevant data about the public networks to provide decision-makers with a holistic view of the situation (see figure 2). This includes knowledge of the supply status per building as well as cumulative statistics per infrastructure. In addition, affected buildings are coloured in order to quickly identify relevant areas.

During the use of the ELD, other software requirements were identified, such as real-time map updates. The implementation of these features is planned for the future. In summary, all results produced in a coupled co-simulation can be integrated into the ELD. However, the data is aggregated and filtered to produce a high-level visualisation and to provide easy access to the data for authorities and decision makers.

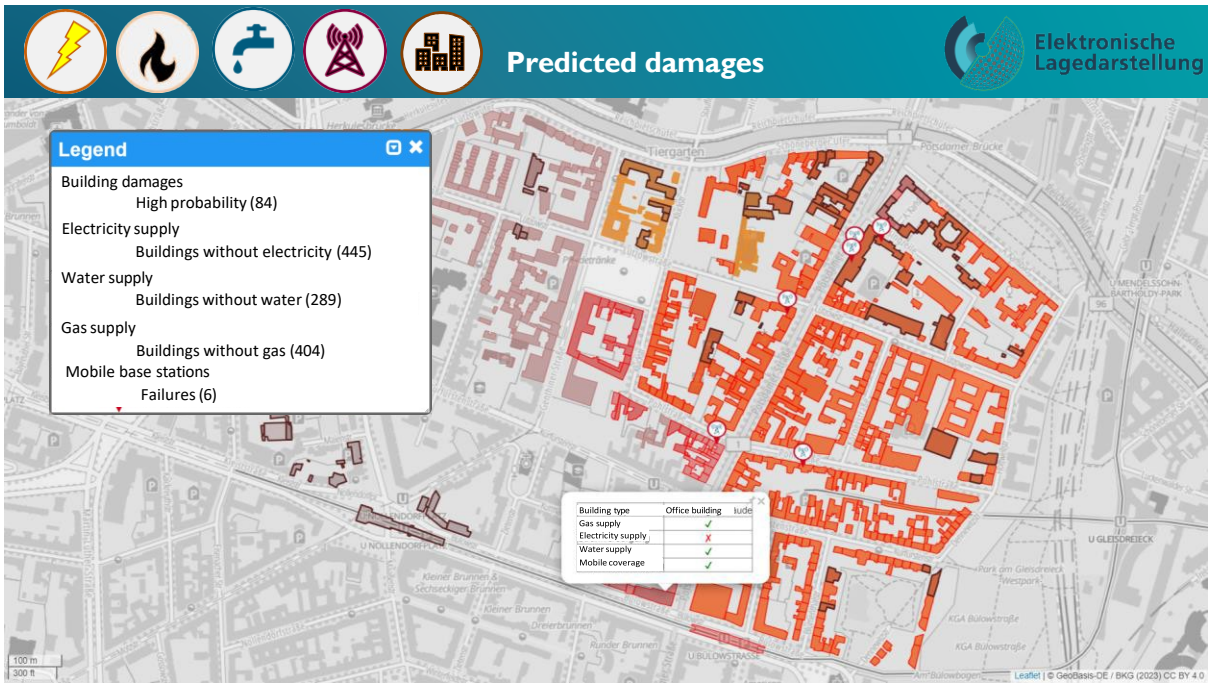


Figure 2. Example visualization in ELD of a predicted multi-layered failure cascade of critical infrastructures (Map data: CC BY 4.0: © GeoBasis-DE / BKG (2023) CC BY 4.0[2])

MobiKat®

Figure 3 shows a simulated crisis scenario in MobiKat® (Fraunhofer IVI MobiKat®, n.d.). It displays the same situation as in figure 2 but focusing on the first responders and respective incidents with subsequent tasks which are partly connected to the infrastructure failures displayed in figure 2. Different types of tactical units from police, fire department and emergency response are assigned to separate missions according to the needs to resolve realistic incidents (see tree view on the left in figure 3). The location and status of each unit is updated dynamically and displayed with its tactical sign on a map in MobiKat® (see map part in figure 3). Additionally, different kinds of statistics characterizing the situation and the availability of resources are being integrated into MobiKat® (see right hand side of figure 3). This provides a good overview of the considered situation and the distribution of resources tailored for users from authorities and organizations responsible for public safety.

MobiKat® is chosen here, because it is a well-established, modular system for the operational command and administration of the corresponding necessary resources. Its user groups are the command staff and relief forces from fire and rescue services, as well as disaster managers and police forces. The system offers various hardware applications for all command levels. The main features are:

- Display of geoinformation about the operation site – e. g., traffic network, aerial images, buildings, land use, information about the terrain
- Visualization of operation-specific information – e. g., flooding zones, tapping points for extinguishing water, critical infrastructures, current water levels, weather, construction sites
- Overview of available staff and resources
- Overview of damages, affected persons and measures
- Documentation of the operation
- Algorithms for the optimization of decision support
- Synchronization of all situational information
- Offline capability and local data storage

In addition to the distributed operational command feature working across levels and organizations with an integrated map, MobiKat® functionalities also include the administration of tasks and appointments, overviews adapted to the current situation as well as analysis and documentation options.

[2] https://creativecommons.org/licenses/by/4.0/

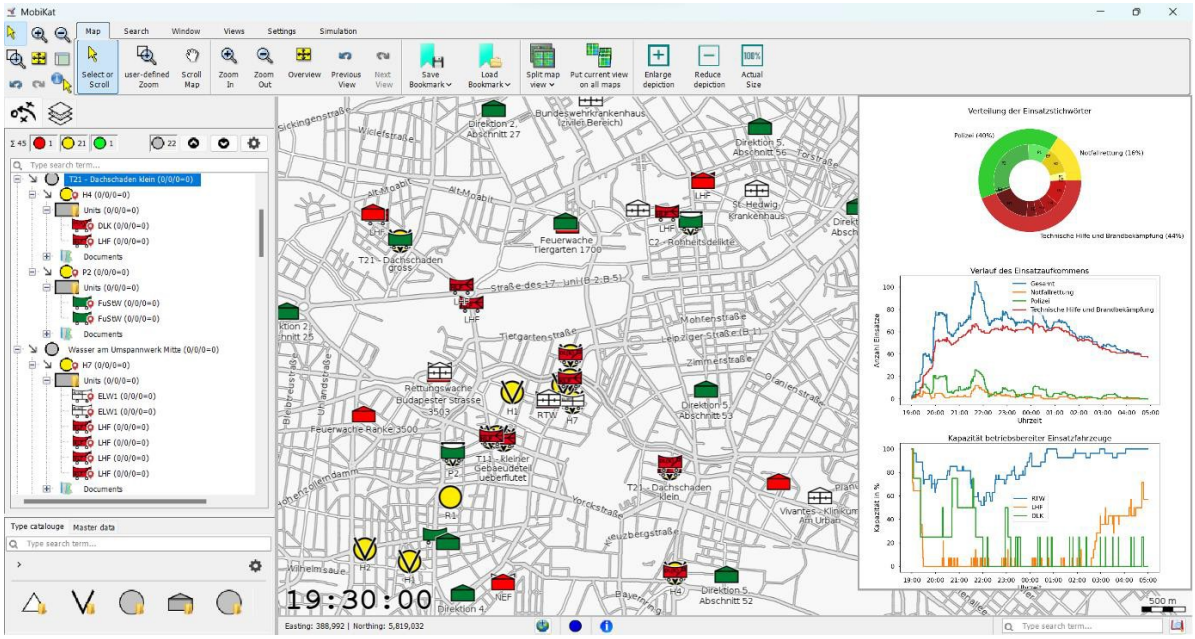


Figure 3. Example visualization of the activity of first responders in MobiKat® (Fraunhofer IVI MobiKat®, n.d.) during a predicted multi-layered failure cascade of critical infrastructures. (Map data: © 2022 OpenStreetMap contributors[3], © GeoBasis-DE / BKG 2019 (data changed))

Summarizing, the visualization and analysis of the aggregated and filtered data of the digital twin in MobiKat® is especially beneficial for users at authorities and organizations responsible for public safety.

DigLT

The *Digitaler Lagetisch* (DigLT, engl. *Digital Map Table*) is a versatile software system designed for collaborative situation visualization and analysis. It allows multiple users to work simultaneously on the same situation using personal computers, digital tables, or other visualization devices (van de Camp et al., 2021; Hoppe et al., 2020). The software is modular and customizable, making it suitable for various applications ranging from education to mission planning, execution, and review. It can integrate diverse data sources and geodata to provide relevant information for each use case, enabling users to make informed decisions.

In addition, the DigLT offers VR capabilities, allowing users to immerse themselves in the situation and experience it in 3D (Strentzsch et al., 2017; Hoppe et al., 2018). Figure 4 shows a screenshot of the VR environment of the DigLT with relief units depicted with tactical symbols. This feature enhances situational awareness and provides a more realistic and immersive visualization of the environment, which is especially useful for training and preparation for future missions. The symbols and icons displayed can be adapted to the needs of the respective user groups (e.g., emergency services, military), which enables a smooth transition from other situation reports.

In DigLT, the concept of individual visual layers allows for easy adjustment and filtering of map information, with the ability to stack and blend the layers for a comprehensive view. The software adheres to open standards, ensuring seamless access to various data sources. For instance, it offers extensive support for OGC and NATO standards, facilitating easy integration of geographical data.

Currently, the digital map table does not have an interface to the *OGC SensorThings API*. However, it is planned to implement this interface in the future. Once implemented, the interface will allow for the integration of the co-simulation results filtered according to the user's needs. For each simulator, this integration will involve creating a new layer on the map that displays all *Things* as corresponding icons or geometries based on the respective

Locations and time stamps. *Datastreams* and *Observations* will be made available as properties associated with each *Thing*. Thus, the DigLT's timeline tool gives the user access to the time-based progress of observed properties, e.g., water pressure or telecommunication network availability at selected sites and the cascading impact any status change has on other infrastructure. As a proof of concept, an interface between the operational

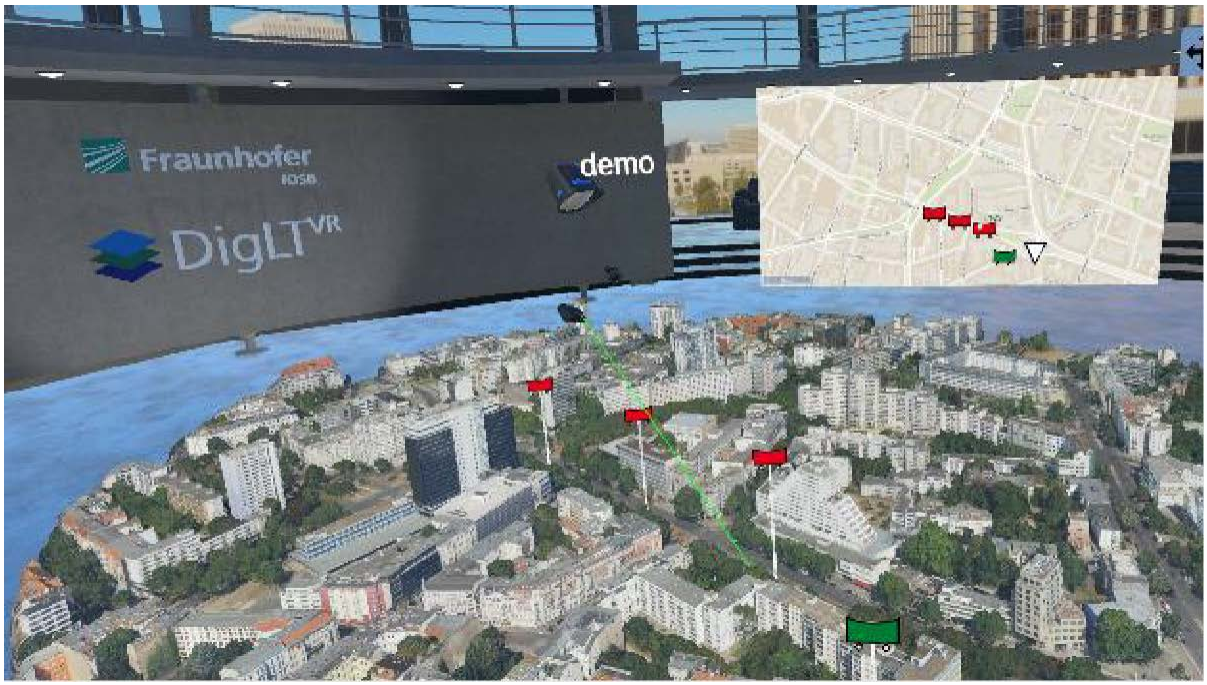


Figure 4. Example visualization of relief units in the virtual reality (VR) view of the digital map table (DigLT). A user is inspecting the current positions of emergency vehicles in the city center of Berlin. (Textures: © 2024 IBCAO, Landsat/Copernicus, Google; Map data: © 2024 GeoBasis-DE / BKG, OpenStreetMap[4])

dynamics simulation module and the DigLT has been established enabling spatial and temporal visualization of incoming emergencies and available relief units both in 2D and VR (see figure 4).

DISCUSSION

A comparison of the three tools for the visualization of simulation and sensor data, ELD, MobiKat®, and DigLT, reveals distinct functionalities that cater to different user groups and operational requirements. Each system offers a range of different functions, aggregates the data in a different way and therefore provides a different form of visualization of a (critical) situation.

On the one hand, the DigLT stands out with its VR capabilities, allowing users to immerse themselves in the situation. The advanced level of this features enables realistic scenarios and enhances situational awareness, both for individuals and in cooperative mode. On the other hand, these VR capabilities require more sophisticated visualization devices and VR equipment. The tool is therefore less suitable for widespread use in critical situations than for mission preparation and training.

The ELD provides decision-makers at all levels of government with a holistic view of the situation. The lack of real-time map updates makes it difficult to visualize dynamic situations and shifts the focus to a more granular view. In addition, the web-based platform offers functions to support staff work and crisis management in terms of communication, information and data exchange.

MobiKat® supports distributed operational use on all command levels and across organizations in desktop, web-based and mobile applications. Compared to ELD and DigLT, it has a more operative and tactical orientated interface aimed at authorities and organization for public safety. At such organizations it is used extensively in daily business as well as crisis situations.

Data integration and usable data sources are robust across all platforms. ELD, MobiKat® and DigLT support open standards, facilitating access to various external data sources. MobiKat®'s modular system allows for a comprehensive overview of public safety-related data. However, the DigLT currently lacks an interface to the OGC SensorThings API, limiting its ability to integrate FROST's sensor data. By addressing this gap in future updates, the DigLT will soon be able to adapt to the presented data processing and visualization pipeline.

In terms of data filtering and preparation for visualization, ELD, MobiKat® and DigLT use a layer-based approach that allows the user to tailor the data displayed. MobiKat® provides an overview of damage, resource and incident-

[4] https://www.openstreetmap.org/copyright

specific data, emphasizing the focus on immediate emergency management.

Recommended user groups for each tool align with these differences. The ELD is well suited for authorities and decision makers at all levels of crisis management who need a comprehensive overview of the situation with spatial data integration. MobiKat® is tailored for executives, emergency responders and disaster managers who need operational command and control and real-time updates. Finally, DigLT is recommended for users who have to deal intensively with the situation, such as emergency services, in order to carry out in-depth analyses and training.
When connected to the digital twin representation through the standardized interface, these tools offer complementary perspectives on the available data gathered in a critical situation and allow for efficient and possibly collaborative work flows in crisis management.

OUTLOOK

Quick action is often required to successfully respond to a critical situation. However, the ever-increasing amount of data available makes it difficult to act quickly in such situations. Data must be aggregated and prepared. A one-size-fits-all approach is not possible as different groups with different needs, such as decision makers, authorities, or infrastructure operators, are involved in the situation. User-tailored aggregation and visualization is therefore crucial for efficient crisis management, and different user groups may require different views of the situation.
The presented visualization approaches access a co-simulation framework (Martini et al., 2024) using the *OGC SensorThings API* standard. This allows for easy integration, also of other data sources, without the need to implement specific interfaces. Other visualization tools addressing further user groups can be easily added.
All presented approaches meet the objective of providing a visualization appropriate to the targeted user group. However, more work is needed to achieve seamless integration of the co-simulation framework and user-tailored visualization. Future work will therefore focus on two main areas: The further development of the tools already in use and the extension of the range of tools.

ACKNOWLEDGMENTS

Part of the authors' research and development, as presented in this paper, is conducted within the Fraunhofer Center for the Security of Socio-Technical Systems (Fraunhofer SIRIOS) funded by the German federal government and the state of Berlin. The authors would like to thank Nils Hückstaedt for his help in creating the visualization within the digital map table and Simon Smaczny for providing background information about the DigLT and proofreading the article.

REFERENCES

van de Camp, F., Hoppe, A. H., Unmüßig, G., Peinsipp-Byma, E. (2020). Cooperative and location-independent terrain assessment for deployment planning using a 3D mixed reality environment. *Proceedings of the SPIE Conference on Virtual, Augmented, and Mixed Reality (XR) Technology for Multi-Domain Operations*, 1142604. doi: 10.1117/12.2558593

Fraunhofer IOSB *FROST*: Fraunhofer Open Source SensorThings API Server, 2023. https://github.com/FraunhoferIOSB/FROST-Server

Fraunhofer IOSB *WebGenesis – System for networking a wide range of data to support decision-making*, 2024. https://www.iosb.fraunhofer.de/en/projects-and-products/webgenesis.html

Fraunhofer IVI *MobiKat*®. (n.d.). Copyright. https://www.mobikat.net/

Hertweck, P., Hellmund, T., van der Schaaf, H., Moßgraber, J., Blume, J.W., Management of Sensor Data with Open Standards, 16th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2019. Conference proceedings, 2019.

Hoppe, A. H., Westerkamp, K., Maier, S., van de Camp, F., & Stiefelhagen, R. (2018). Multi-user Collaboration on Complex Data in Virtual and Augmented Reality. *Proceedings of HCI International 2018. Communications in Computer and Information Science*, 851, 258–265, Springer. doi: 10.1007/978-3-319-92279-9_35

Hoppe, A. H., van de Camp, F., & Stiefelhagen, R. (2021). ShiSha: Enabling Shared Perspective With Face-to-Face Collaboration Using Redirected Avatars in Virtual Reality. *Proceedings of the 4th ACM Conference on Human-Computer Interaction*, CSCW3, 251, ACM. doi: 10.1145/3432950

Liang, S.H., & Khalafbeigi, T., van der Schaaf, H., OGC SensorThings API Part 1: Sensing Version 1.1. OGC® Implementation Standard. 2021.

Martini, T., Rosin, J., Vetter, J. Z., Neuhäuser, S., Lukau, E., Catal, F., Boigk, M., Simon, M., Monteforte, M., Gerold, M., Phung, W., Dietze, S., Finger, J., Brausewetter, P., & Nicolai, S. (2024). Towards a modular co-simulation framework for the assessment of cascading effects among critical infrastructures and the impact on citizens. To be published in the forthcoming proceedings of the *63rd ESReDA Seminar: Resilience assessment - Methodological challenges and applications to critical infrastructures, hosted by the Joint Research Center (JRC), Ispra, Italy, October 25 – 26, 2023*

Rehak, D., Markuci, J., Hromada, M., & Barcova, K. (2016). Quantitative evaluation of the synergistic effects of failures in a critical infrastructure system. *International Journal of Critical Infrastructure Protection*, *14*, 3–17. doi: 10.1016/j.ijcip.2016.06.002

Rehak, D., & Hromada, M. (2018). Failures in a critical infrastructure system. In *System of System Failures*. doi: 10.5772/intechopen.70446

Russell, L., Goubran, R., Kwamena, F., & Knoefel, F. (2018). Agile IoT for critical infrastructure resilience: Cross-modal sensing as part of a situational awareness approach. *IEEE Internet of Things Journal*, *5*(6), 4454–4465. doi:10.1109/jiot.2018.2818113

Strentzsch, G., van de Camp, F., & Stiefelhagen, R. (2017). Digital Map Table VR: Bringing an Interactive System to Virtual Reality. *Proceedings of the 15th International Conference on Virtual, Augmented and Mixed Reality. VAMR 2017*. Lecture Notes in Computer Science, 10280, 54–71, Springer. doi: 10.1007/978-3-319-57987-0_5

van Rees, Eric (2013). Open geospatial consortium (OGC). Geoinformatics, 16, 28-29

Strategies for Crisis-Responsive Governance: Automated Anomaly Identification in Public Services

Hakan Unveren
University of Vermont
hunveren@uvm.edu

Ayesh Abu Lehyeh
University of Vermont
ayesh.abulehyeh@uvm.edu

Duygu Pamukcu
HEC Montréal
duygu.pamukcu@hec.ca

Christopher W. Zobel
Virginia Tech
czobel@vt.edu

ABSTRACT

This paper introduces a machine learning tool for service systems, focusing on accurate classification of service requests and swift anomaly detection, particularly crucial during emergencies. Employing a Support Vector Machine model, this tool automatically classifies service calls into predefined categories with high accuracy, while effectively detecting irregular requests that require specific attention from operators. This approach streamlines resource management by reducing the manual categorization workload and enables early detection of emerging service needs. Examining Orange County, Florida 311 System data, with a specific focus on the COVID-19 period, we illustrate the tool's success in automatic request categorization and anomaly detection. Overall, this tool presents an effective automation approach to help with efficient resource management of service systems and proactive assessment of public service needs, promising to revolutionize service request management during crises. Future work will explore additional classification models for enhanced accuracy and integrate automated alerts for proactive disaster management.

Keywords

311 services, classification, anomaly detection, supervised learning, disaster management.

INTRODUCTION

The evolution of information technology-enabled service systems has revolutionized the way governments interact with their constituents. Public service requests offer individuals a direct means to voice their needs and concerns, serving as a vital bridge for citizen-government interactions (Schwester et al., 2009; Wu, 2020). Citizens now have the flexibility to engage with government agencies through various mediums such as online portals, mobile applications, and chat systems. These systems facilitate the seamless collection of requests and enable prompt and effective responses, thereby fostering a more responsive and citizen -centric governance approach. Given the widespread availability and accessibility of public systems, the necessity for streamlined and automated processes becomes imperative to meet the diverse and expanding needs of the population (Hashemi, 2022). Consequently, the automation of handling service requests in public systems emerges as a crucial step toward enhancing efficiency and responsiveness in governance.

Automation in public services refers to the integration of technology and streamlined processes to carry out tasks and functions within government agencies without direct human intervention. This involves the use of software, artificial intelligence, and other technological solutions to handle routine and repetitive tasks, facilitating the analysis of patterns in service requests, anticipation of citizen needs, and optimization of resource allocation

within government agencies (Bianchini et al., 2006; Brogi et al., 2006; Cheng et al., 2022; Pamukcu et al., 2022). The overarching goal is to improve the efficiency of public service delivery and enable agencies to proactively address emerging challenges. The benefits of automation extend to quicker response times, a consistent service approach, and the reduction of human errors, fostering a reliable interaction between citizens and government agencies. Moreover, as automation takes charge of routine tasks, it frees up human agents to concentrate on more complex and critical issues that demand a personalized touch.

The importance of automation becomes even more evident during disasters (Shah et al., 2019). Community members often actively contribute to reporting emerging needs during crises, and automated systems can leverage crowdsourced data to comprehensively identify and categorize new service requirements, providing a more nuanced understanding of the evolving situation (Pamukcu et al., 2021). Accurate categorization of emerging needs and efficient resource utilization becomes imperative to ensure a dynamic and responsive system, even during surges in service requests. This heightened demand for accuracy and efficiency underscores the crucial role that automation can play in effectively managing and responding to the diverse and urgent needs that arise during disaster scenarios.

The recent coronavirus pandemic highlighted critical challenges in current public service request management practices as systems faced a surge in pandemic-related calls, particularly in the early months of the emergency (Filip et al., 2022). The overwhelming volume of calls resulted in an increased demand for system resources, including service operators and personnel hours (Pamukcu et al., 2021). Manual operations, such as assigning each request into its appropriate category, consumed a substantial amount of limited system resources. Additionally, systems began to receive irregular types of requests that did not align with existing categories, posing request categorization challenges, and potentially causing delays in response times as these inquiries required more careful attention.

An illustrative example is seen in the Orange County Florida 311 non-emergency system, where similar requests were inconsistently assigned to different service categories. To address this issue, the system created new categories to address newly emerged service needs under a unified classification. Despite early signs of the pandemic, with the first two cases in Florida reported on March 1 and the declaration of a state of emergency on March 9, a dedicated category for COVID-19 was not established until May 8, indicating a notable delay in adapting to emergent situations (Cutway, 2020). Similarly, the New York City 311 system introduced new categories such as “Social Distancing” and “Mask Covering” to accommodate pandemic -related concerns, highlighting the necessity for a more adaptive categorization process within these systems (Pamukcu et al., 2021). The examples highlight the ongoing need for refined and efficient automation processes, particularly in times of crises.

Automated systems, equipped with adaptive algorithms, effectively tackle challenges arising from the extensive need for human resources in manual operations. This adaptive capability ensures the system remains dynamic and responsive to the evolving nature of service requests. For example, the literature indicates that the incorporation of Natural Language Processing (NLP) in automated systems enhances their ability to comprehend and categorize unstructured data, such as text or voice messages, from affected individuals. Additionally, numerous studies emphasize that the application of machine learning and neural network methods can provide accurate service categorization (Bianchini et al., 2006; Brogi et al., 2006; Hashemi, 2022). This is particularly valuable in identifying newly arising type of service calls that may not be predefined in the system and appear as irregular requests or anomalies. Utilizing such models, systems can become capable of recognizing and categorizing these anomalies based on distinctive patterns and characteristics.

To tackle the challenges associated with categorizing service requests, particularly during crisis periods, our research employs automation techniques to introduce an innovative tool that integrates classification with anomaly detection. This tool automatically categorizes regular requests while promptly identifying anomalies, aiming to improve the adaptability and effectiveness of service management. The goal is to ensure that systems can respond efficiently to unprecedented events, thereby better serving the community in times of crisis.

In this study, we specifically investigate the 311 system of Orange County, Florida, to demonstrate the potential improvements in handling service requests through accurate classification and simultaneous anomaly detection, especially during disaster scenarios. Our analysis is supported by a comprehensive dataset of service requests obtained through collaboration with the Orange County 311 government. This partnership has afforded us access to valuable data, facilitating a thorough evaluation of the system's performance and the potential impacts of our proposed tool.

This study makes a valuable contribution to the efficient management of service systems, especially during crises. First, we introduce a novel and practical method for automatic request classification with high accuracy. We employ Support Vector Machines (SVM), a supervised machine learning technique, to classify service calls. The choice of SVM is motivated by its notable advantages, including rapid processing and robust performance, even with a relatively small dataset. The model is particularly beneficial for text classification tasks, often constrained by the availability of limited labeled samples. Consequently, our approach offers a practical and accurate solution for service providers, empowering them to enhance their operations through automated categorization.

Second, our method innovatively utilizes the decision function of SVM to incorporate anomaly detection capabilities. By adding an "extra category" to the predefined set, defined as "Other," our model dynamically adjusts to unfamiliar types of calls. This adaptability ensures the system remains flexible and responsive to evolving service needs. Moreover, this enhancement significantly reduces the time needed to detect new categories, a crucial step in improving efficient service management, particularly during crises. This feature substantially improves the system's responsiveness and efficiency during crises, ensuring timely and appropriate service delivery.

RELATED WORK

In the literature, there is a notable gap in addressing challenges related to service request management, particularly with a focus on efficient and adaptive service categorization, especially during disasters. This summary encapsulates a few studies utilizing machine learning and artificial intelligence techniques on service data and relevant efforts in classification of crisis-related social media data to help with request management efforts of service systems.

Several studies focus on identifying service request patterns, volumes, and frequencies in different regions. For example, Cheng et al. (2022) employ machine learning techniques to forecast the number of non-emergency service requests in the Miami-Dade County and to identify the neighborhood traits within census tracts that influence call volumes, while Kontokosta et al. (2017) use gradient boosting regression to classify heat and hot water problems in New York City. Wang et al. (2017) explore the use of non-emergency service requests for decision support and socioeconomic trend prediction, employing k-means clustering to categorize census tracts.

Hashemi (2022) addresses automatic detection of service request types using recurrent neural networks, primarily focusing on classifying requests into preexisting categories. While our paper shares similarities with Hashemi (2022), our approach distinguishes itself by incorporating timely detection of anomalies, which demands more careful attention from service providers. Brogi et al. (2006) and Bianchini et al. (2006) discuss the use of semantic descriptions and ontologies in service discovery. Specifically, Brogi et al. (2006) present a matchmaking system for service compositions and Bianchini et al. (2006) propose a semantic-enriched framework for service description and discovery. Unlike previous work, our tool is designed to simultaneously identify outlier service requests, i.e., anomalies, that do not fit within any of the existing service categories while conducting automatic classification of requests.

In addition to efforts for classifying requests to service systems directly, recent studies have also focused on utilizing social media data for service providers' attention. With the increasing use of social media during disasters, relevant information extracted from such platforms has been found helpful for the crisis management efforts of local authorities. Li et al. (2018) utilized SVM among other techniques to classify tweets with high accuracy, but their approach lacks the anomaly detection concept that is helpful for service systems to differentiate the newly arising needs during crises from regular service needs. Similarly, Alam et al. (2019), Li et al. (2021), and Wiegmann et al. (2020) focused on classifying disaster-related content in social media data without the capability of detecting unusual disaster types.

Kersten et al. (2021) propose a machine learning model which improved the detection rate of crisis-related tweets, but this came with a higher false positive rate, leading to more non-crisis tweets being misclassified as crisis-related. Other studies, such as Aipe et al. (2018) and Buntain et al. (2022) were limited to classifying known crises and lacked the detection of unforeseen events. However, our tool extends the SVM model with a distinctive learning approach to identify unforeseen crisis-related requests like those specific to the COVID-19 pandemic, without prior training or information on such specific scenarios. This enables our model to adapt dynamically to emerging crises and enhance its effectiveness in crisis management.

In summary, while previous literature underscores the potential of automated techniques, semantic descriptions,

and ontologies in service requests and relevant social media content, there is limited exploration of machine learning approaches for disaster management efforts in service systems. Our study aims to fill this gap by proposing a dynamic model for accurate service request classification and timely anomaly detection . By doing so, our tool contributes to the advancement of knowledge in this domain, especially in adaptive service management during emergencies. This distinctive approach sets our study apart from existing research, offering valuable insights into automation and efficient decision-making.

PROBLEM CONTEXT: 311 SERVICE SYSTEMS

Cities have increasingly embraced 311 systems for local non-emergency service management, driven by two key factors (Samuel, 2019). First, the implementation of 311 services aims to alleviate the workload on 911 operators by diverting non-emergency inquiries to dedicated 311 call centers. In doing so, calls related to non-emergency issues can be efficiently managed by operators with expertise in directing such concerns to the appropriate municipal departments. Secondly, these services serve as a convenient channel for residents to request serv ice and access information on a diverse range of municipal topics, spanning from garbage collection and snow removal to tax regulations and new local laws. This enhances citizen engagement and provides a centralized resource for obtaining relevant information. Consequently, 311 services often encompass a variety of predetermined categories for which residents might seek assistance. This structured approach ensures that residents receive prompt and specialized attention for their specific concerns, contributing to more effective municipal service delivery (O'Brien, 2016).

Over 300 cities and counties in North America have integrated 311 as a fundamental resource for municipal information and services (Winbourne, 2022). These systems function independently from each other, leading to a lack of uniformity in their operational frameworks. Different cities adopt various strategies in their service management based on their local needs, preferences, and agreements with business partners. For instance, while Chicago 311 fully incorporates automation technologies to streamline and categorize service requests(City of Chicago, 2018). Orange County Florida 311 favors manual categorization of requests and relies on human operators for the management of calls, showcasing the diverse approaches cities take in utilizing 311 services (Pamukcu et al., 2021). The significance of these systems in helping with local disaster management efforts was particularly highlighted during the COVID-19 pandemic when 311 served as a vital conduit for disseminating information critical to safeguarding the health and well-being of residents and their families (Pamukcu et al., 2021).

DATA AND METHODOLOGY

Data Description

For our study, we examine the Orange County, Florida 311 system to demonstrate accurate classification and prompt anomaly detection. Our analysis is supported by comprehensive service requests data from OCFL311, obtained through collaboration with the Orange County 311 managers. The dataset comprises 1,055,678 records spanning 2016 to 2021 with 14 predefined main categories of service calls. A record in the context of this study refers to an inquiry, which is created upon the receival of the service request. Each record is uniquely stored in the database, and it undergoes updates as necessary, such as follow-ups or closure upon resolution or forwarding to a specific municipal depa rtment. The deliberate choice of the Orange County 311 system was driven by its commitment to meticulous manual processing of service requests, with a primary focus on achieving the highest level of accuracy in request categorization.

Upon receiving service requests through various channels such as phone calls, chat, or mobile apps, each entry undergoes a comprehensive triage and processing phase facilitated by trained 311 operators. This careful handling is designed to guarantee the precise assignment of requests to their respective and appropriate categories. It's important to note that the manual triage of requests places a substantial demand on system resources, and the efficiency of this process directly impacts the overall performa nce of the system, as highlighted by Pamukcu et al. (2022). This underscores the significance of optimizing the request categorization process to enhance system efficiency and effectiveness.

The Orange County Florida 311 dataset offers comprehensive information about each request, including the exact location of the incident, detailed issue descriptions, along with the issue creation and completion dates, and resolution information. To enhance manageability and focus on essential aspects relevant to our study, we only retained critical information pertinent to our study, such as the main categories and detailed subcategories of the requests, along with descriptions and creation dates of the reported issues. These retained columns are instrumental for training our model, providing valuable insights into the nature and classification of service

requests. The subsequent section will delve into the specifics of our data preprocessing methodology, providing a clear understanding of the steps taken to refine and prepare the dataset for our analytical processes.

Data Preprocessing

In this stage, we meticulously conducted data preprocessing as an integral part of our approach to text classification. Recognizing the importance of clean and well-structured data for the success of machine learning models, we employed a series of preprocessing steps. This included the removal of punctuation, conversion to lowercase, and elimination of stop words to refine the textual information and enhance its focus. We also addressed missing data to ensure the dataset's reliability and consistency. Tokenization and vectorization were applied to transform the text into a format suitable for our chosen machine learning model.

In order to transform text into a format suitable for our chosen machine learning model, we opted for the widely used Term Frequency-Inverse Document Frequency (TF-IDF) vectorization method. TF-IDF proved to be a robust choice, capturing the importance of terms in the context of the entire dataset, and providing a numerical representation that retained the significance of words while mitigating the impact of common term s (Shah et al., 2019). Furthermore, we labeled and encoded categorical labels, facilitating numerical representation. Our commitment to thorough data preprocessing played a crucial role in preventing overfitting and promoting efficient model training. The resulting clean and well-structured data contributed to the interpretability of our model results and facilitated the identification and resolution of potential issues.

Automatic Classification of Issue Descriptions with SVM

SVM is a supervised machine learning model that utilizes algorithms for classification problems (Cortes & Vapnik, 1995). The SVM model is trained on a set of labeled training data for each category, enabling it to categorize new text. In our investigation, we leveraged the SVM algorithm for the intricate task of multi-class classification in the context of issue description categorization. Originally designed for binary classification, SVM is extended to handle multiple classes through a one-versus-the-rest (OvR) approach (Cortes & Vapnik, 1995; Hsu & Lin, 2002). The OvR involves training a binary classifier for each class against the rest, and the class with the highest confidence score is assigned during prediction (Galar et al., 2011).

Specifically, the classification task was approached using the Linear Support Vector Classification (LinearSVC) algorithm, a variant of SVM tailored for high-dimensional datasets (Gualtieri & Cromp, 1999). The core hyperparameter governing the model's behavior is the regularization parameter. This parameter, set by default to 1.0, serves as the inverse of the regularization strength. Higher values of this parameter lead to reduced regularization, allowing the model to fit the training data more closely. Striking the right balance for this parameter is crucial as it determines how much emphasis the model places on fitting the training data versus preventing overfitting (Hastie et al., 2009).

The optimization process to prevent overfitting in LinearSVC involves minimizing a cost function that combines a hinge loss term for classification accuracy with a regularization term (Hastie et al., 2009). The Pegasos algorithm, a variant of stochastic gradient descent, is used to iteratively update the model parameters based on a subset of training examples, allowing the model to adapt to the complexities of the data (Shalev-Shwartz et al., 2007) The convergence of the training process was guided by adjusting the tolerance parameter. As the algorithm iterated over the dataset, the training process would stop when the change in the cost function (a combination of hinge loss and regularization) dropped below the specified tolerance. This ensured that the model achieved a satisfactory level of convergence, balancing accuracy on the training data with generalization to new, unseen examples (Apidopoulos et al., 2022).

Anomaly Detection for New Categories

In addition to our primary classification objective, our study emphasizes the early detection of new patterns through the integration of anomaly detection. This facet adds a crucial layer to our automatic classification system, enabling the timely identification of instances that deviated from established categories. The importance of early detection lies in the system's ability to proactively recognize emerging trends, ensuring a responsive approach to evolving data dynamics.

Anomaly detection, as a pivotal component, operated by identifying outliers within the dataset —instances that exhibited novel or unexpected patterns. These outliers, which are a new type of call, were promptly categorized as “Other” in a dynamic and iterative manner, marking a departure from traditional classification frameworks. The iterative nature of this process allowed our model to continuously refine the “Other” category as the dataset evolved, making it a proactive mechanism for adapting to shifts in patterns and trends.

In the technical intricacies of our automatic classification system, the decision function scores play a crucial role, offering a nuanced and data-driven approach to categorizing instances beyond predefined labels. The decision function scores, obtained through the LinearSVC algorithm, provide a measure of the distance of each instance from the decision boundary. These scores represent the model's confidence in its classification decision, with lower scores indicating proximity to the boundary.

The importance of finding the optimal decision score threshold cannot be overstated. This threshold delineates the boundary between instances classified within established categories and those categorized as "Other". The process of determining the threshold involved careful calibration to strike the right balance between sensitivity and specificity in anomaly detection. Sensitivity ensured the accurate identification of true anomalies, while specificity prevented the misclassification of instances merely representing variations within known categories.

In the quest for the best decision score threshold, we engaged in a systematic exploration, considering various values to assess their impact on the model's performance. This involved a trade-off analysis, as adjusting the threshold influenced the number of instances classified as "Other" and the potential impact on overall model accuracy.

The technical complexity of this process lies in its adaptability to changing data dynamics. As the dataset evolved, the decision score threshold underwent recalibration to accommodate new patterns and maintain the efficiency of anomaly detection. This dynamic adjustment allowed our automatic classification system to stay attuned to the subtle shifts in data patterns over time.

ANALYSIS AND RESULTS

As previously mentioned, we demonstrate the effectiveness of our approach in automatic request categorization and anomaly detection using the Orange County, Florida 311 System data. The system classifies requests into 14 main categories, as illustrated in Figure 1. The graph indicates the number of service requests and the proportional distribution of main service categories for the period between 2016 and 2021 after preprocessing the data . Notably, the Animals category experiences the highest call volumes, followed by Public Safety, and then Neighborhoods & Housing. This visualization provides insights into the relative frequency of service requests across different main categories during the specified time frame. Figure 1 also shows the imbalance across the 14 main categories, a challenging problem in the realm of text classification. Following data cleaning and preprocessing, which involved the removal of missing and incorrectly labeled data, the number of observations decreased from 1,055,678 to 600,362. This is a crucial step in the text classification before training the machine learning model.

Pandemic-related Requests Categorization in Early Phase

Our analysis is rooted from early signals indicating challenges in categorizing emerging service needs during the initial phases of the COVID-19 pandemic. Following the declaration of a state of emergency on March 9, the Orange County, Florida 311 system started to receive many pandemic-related calls. However, a dedicated category for COVID-19 was not established until May 8, 2020. Following this, the system introduced 23 new subcategories to consolidate pandemic-related inquiries under dedicated classifica tions.

In our preliminary analysis, we observe that similar requests were inconsistently assigned to different service categories. The absence of a specific category for this emergency led to the classification of pandemic -related inquiries into various existing categories, each managed by different departments. To identify the categories to which these inquiries were assigned, we filtered 311 requests based on the issue description attribute using keywords such as "covid," "corona," "pandemic," and "epidemic" directly related to the COVID-19 crisis. Our analysis of this subset revealed that the system had already received 3440 unique pandemic -related service requests, constituting 9.5% of the total requests, from the beginning of the pandemic until the creation of a new category to accommodate calls that did not fit well under any existing category. Similar types of requests were addressed under different main categories, specifically *Public Safety*, *Neighborhoods & Housing*, *Families*, *Health & Social Services*, and *General*.

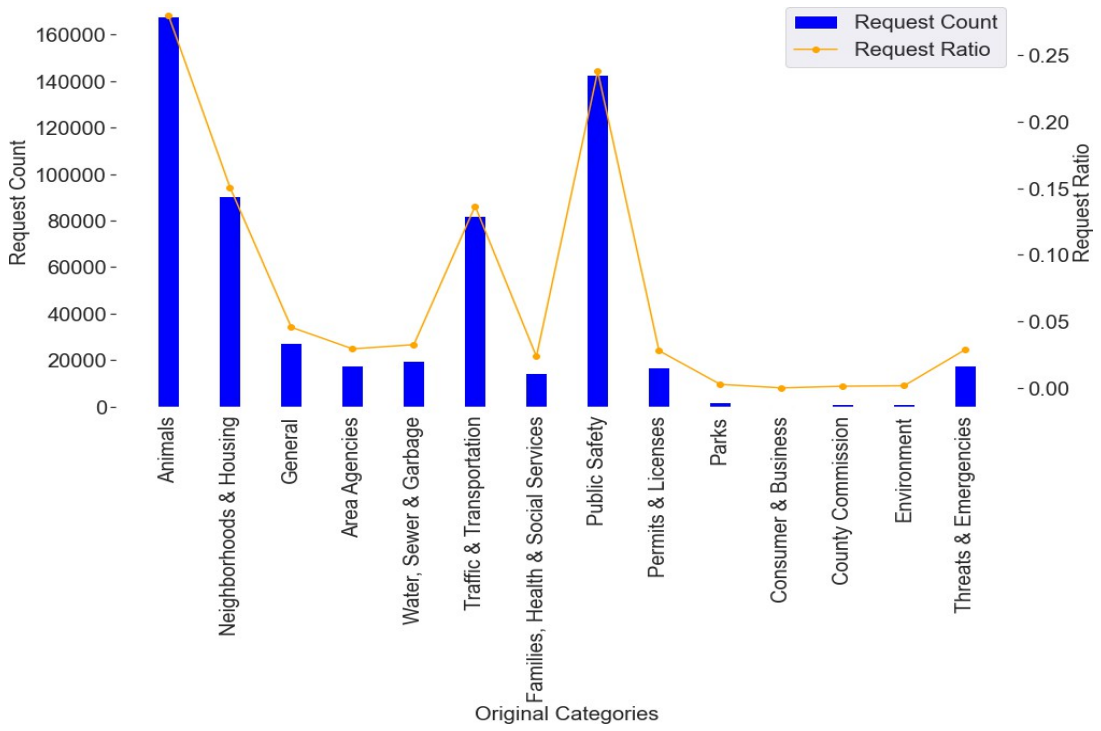


Figure 1. The proportional distribution of main service requests from 2016 until 2021

Support Vector Machine Model Results

The full preprocessed dataset was used in this stage to train and test the data with the LinearSVC algorithm. The dataset first has been divided into training and testing sets where 80% formed the training set and the remaining 20% formed the testing set. The TF-IDF vectorizer is then used to extract features from the data and finally the transformed features are used to fit our model. The testing set is then used for prediction and evaluation of the performance of our classifier model.

Table 1. SVM Model Results

Performance Metric	Result
Accuracy	92.19%
Precision	91.95%
Recall	92.19%
F1 Score	91.96%

Table 1 summarizes the results on the testing set, showcasing the proficiency of the SVM model in classifying service calls. With an accuracy of 92.19%, the model demonstrates remarkable reliability, correctly categorizing service calls with a high degree of certainty. Precision and recall, standing at 91.95% and 92.19% respectively, further underscore the model's precision in minimizing false positives and its comprehensive ability to capture relevant instances. The harmonized balance between precision and recall is encapsulated in an F1 score of 91.96%, highlighting the model's adeptness in maintaining equilibrium between exactness and completeness. Such robust performance metrics not only affirm the model's effectiveness in navigating classification challenges but also bolster its potential as a pivotal tool in the context of service request classification.

In order to thoroughly evaluate the performance of the proposed classification algorithm, we delved into a granular analysis of the model's predictions and the corresponding actual outcomes using the confusion matrix. Figure 2 provides a comprehensive breakdown, offering insights into the algorithm's performance across different true categories and their respective predictions. Within the confusion matrix, the diagonal elements stand as representations of instances correctly classified for each class. The varying shades along the diagonal convey the degree of success in true classification, with darker blue denoting higher accuracy. The figure showcases the

model's robust performance across the majority of categories, except for *General* and *Water, Sewer & Garbage*. In these specific areas, the model's performance appears to exhibit some challenges, suggesting opportunities for further investigation and potential refinement to optimize classification accuracy.

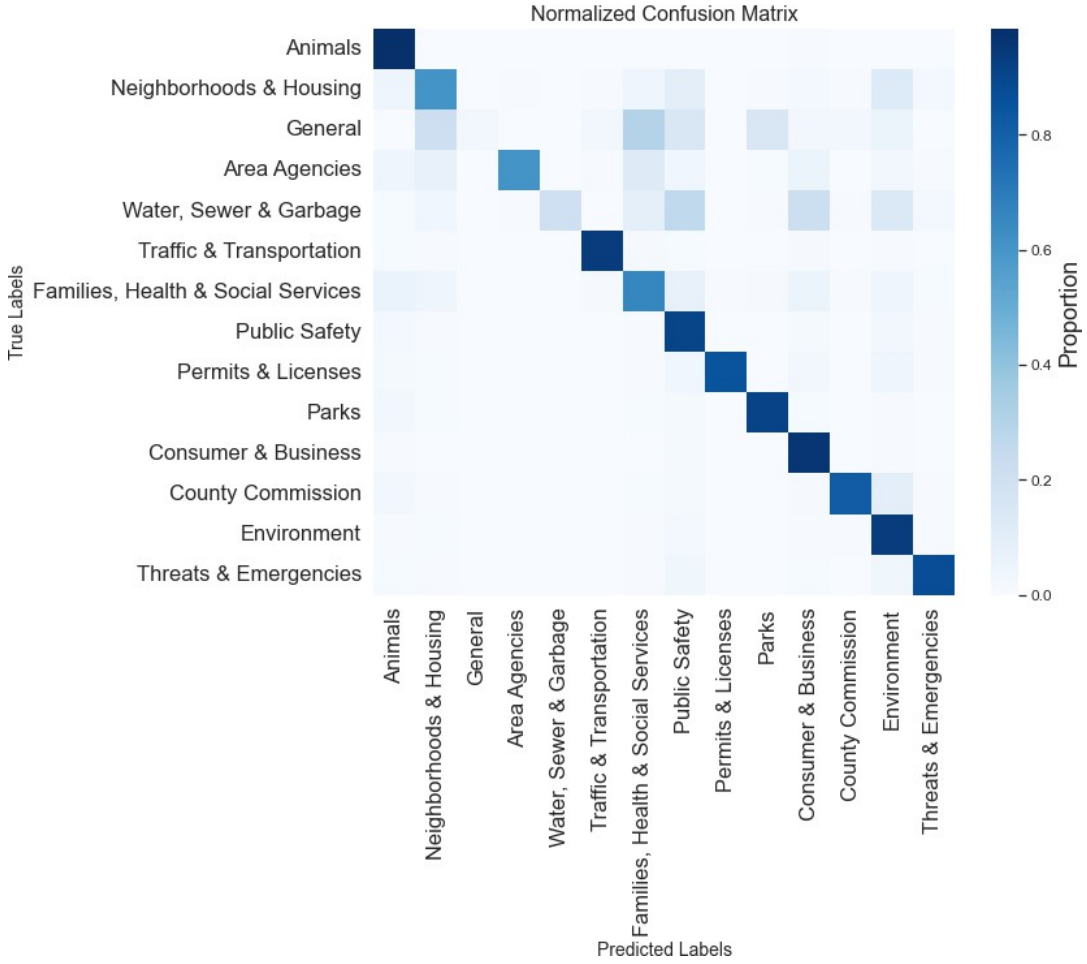


Figure 2. Confusion matrix for multiclass classification of Orange County, Florida 311 service categories

The complexity in classifying requests in the *General* category aligns with its nature, encompassing diverse requests related to information systems, incomplete calls, disconnected calls, undeliverable requests, or general inquiries not directly tied to any other category. Conversely, an examination of request s directed to the *Water, Sewer & Garbage* category reveals shared issues with several other categories, including concerns about water quality, hazardous waste, and standing water. This shared nature poses challenges for the model in accurately classifying these types of service requests. These insights contribute to a comprehensive understanding of the algorithm's strengths and areas for improvement, guiding future efforts to refine and optimize its perf ormance across diverse service request categories.

Anomaly Detection for Unveiling Emerging Categories

In this part of our study, we employed the previously outlined anomaly detection methodology to identify irregularities within the service request patterns. This approach helps recognize unusual request types, signaling for new public needs and a possible emergence for additional request categories. To detect anomalies in service requests, the SVM model was trained on data spanning from January 1, 2016, until December 31, 2019. This timeframe was deliberately selected to encapsulate the anomalies that surfaced during the first year of the pandemic crisis, which broke out in Orange County, Florida on March 9, 2020.

Figures 3, 4, and 5 highlight these irregularities, starting from January 1, 2020 , classified under “Other” category. The analysis reveals a notable number of anomalies in the service calls within the specified time frame. The surge in request volumes in this category immediately following the pandemic outbreak is indicative of atypical request patterns during the crisis. It serves as a clear signal for service operators to consider establishing new service

categories to unify service requests related to the ongoing crisis. Illustrated by the COVID-19 crisis, the presented tool proves valuable in identifying an increase in anomalies within service requests, regardless of the specific emergency scenario.

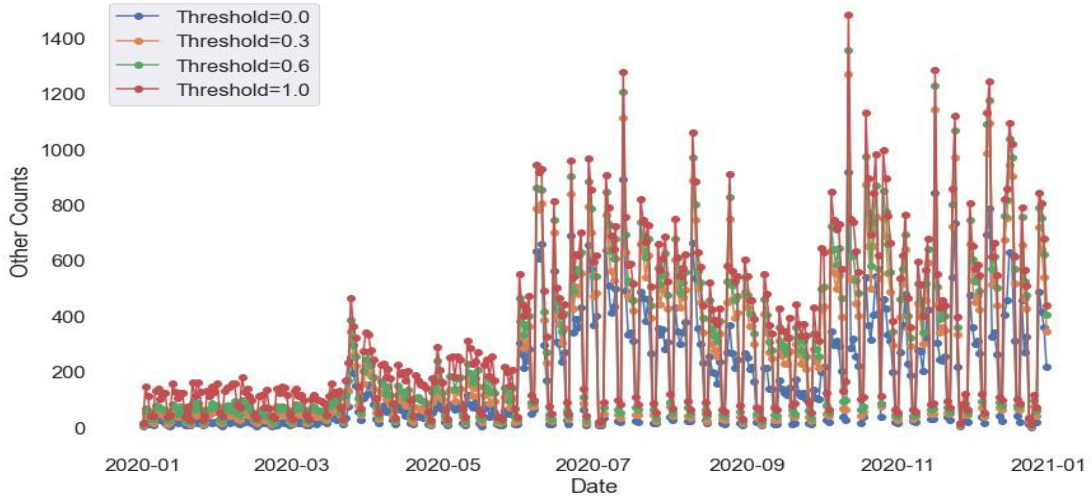


Figure 3. Number of requests classified in "Other" category with different selection thresholds.

To establish the optimal confidence threshold for classification for the "Other" category, we have examined four different thresholds ranging from 0.0 to 1.0, illustrated in Figure 3. In our anomaly detection model, the proximity of a request to each category is expressed by a confidence value, and a threshold is defined to assign requests to "Other" category if the confidence falls below the minimum threshold. This is crucial for capturing an anomaly when a request does not fit well into any of the existing categories, especially when the ratio of "Other" calls to the total number of calls is minimal.

Through a sensitivity analysis on various threshold values, we concluded that a confidence threshold of 0.0 establishes the optimal boundary. While the anomaly pattern remained consistent across different thresholds, our observations indicated that choosing higher thresholds resulted in the model incorrectly classifying requests under the "Other" category since the model became less sensitive. The selected threshold of 0.0 enhanced the accuracy of our model and effectively distinguished between well-known categories and anomalies, ensuring appropriate classification under the "Other" category.

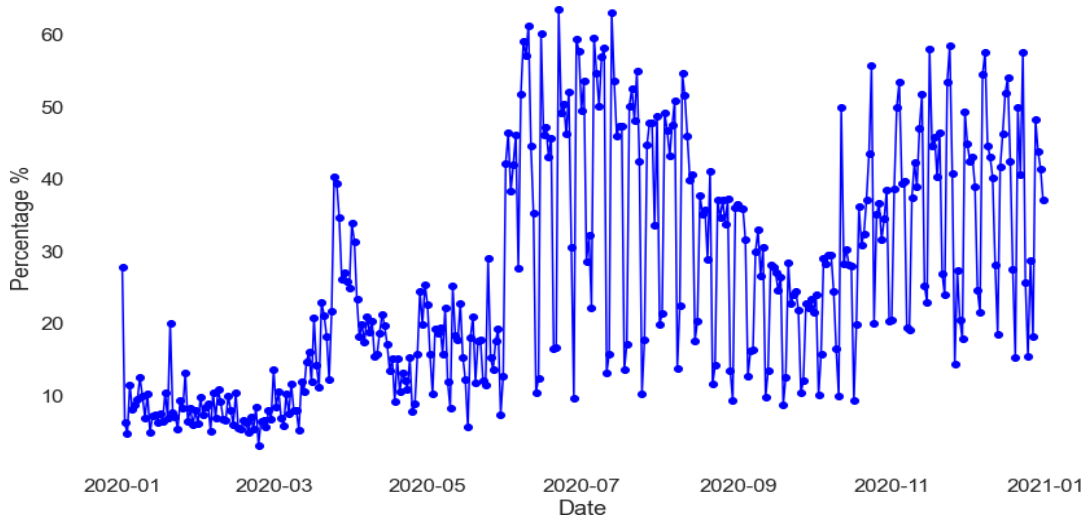


Figure 4. The percentage of requests in the "Other" category among the total number of service requests

Figure 4 illustrates the proportion of "Other" category requests in relation to the overall service inquiries. The data, trained up until January 1, 2020, shows a consistent pattern up to March 2020. Post-March 2020, however, there is a noticeable increase in variability, indicating a rise in service requests falling under the "Other" category.

This period coincides with the intense phase of the COVID-19 pandemic, leading to a surge in unique inquiries in the system. Consequently, the model began to assign these COVID-19 related inquiries to the "Other" category. This finding highlights the utility of our tool in identifying significant increase in unique calls, underscoring a potential need for introducing new categories within the service request system.

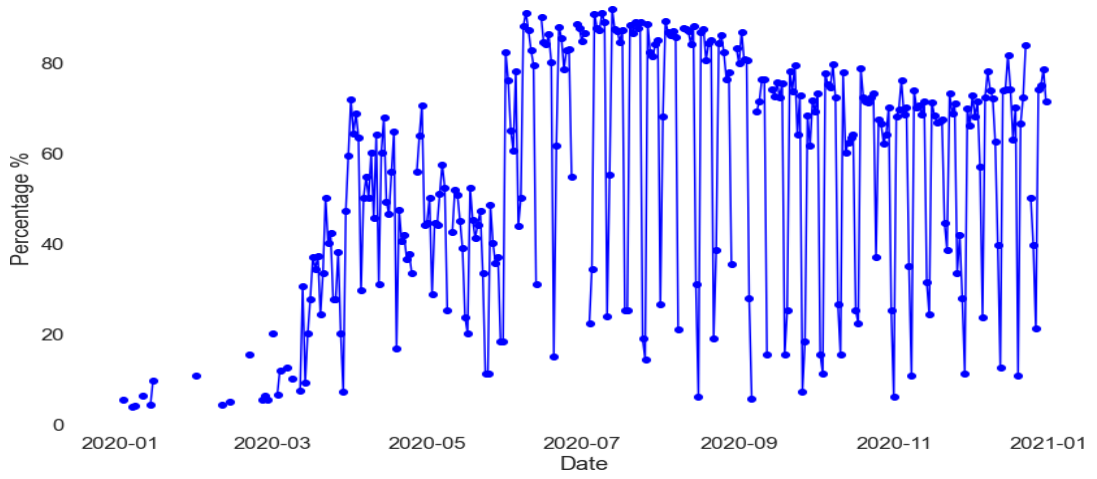


Figure 5. The percentage of pandemic-related anomalies within the total number of requests in the "Other" category

Moreover, to illustrate the proportion of pandemic-specific inquiries within the detected anomalies, we present the percentage of pandemic-related anomalies relative to the total number of requests in the "Other" category. Figure 5 indicates an increase in this percentage after March 2020, aligning with the emergence of new types of calls in response to the state of emergency declared in Orange County, Florida , on March 9, due to the COVID-19 outbreak. Another significant rise is observed in June, coinciding with the announcement of the Coronavirus Aid, Relief, and Economic Security (CARES) Act by the Orange County Government to provide financial support to residents, business owners, nonprofits, municipal and community partners in response to the pandemic (*Orange County Government, Florida*, n.d.)

Consequently, these calls were classified under "Other," contributing to the overall rise in the number of "Other" service requests. These results affirm the effectiveness of our model in identifying such anomalies, which were later addressed by adding new service categories specific to these needs. Specifically, during the COVID-19 crisis, the Orange County Florida 311 system established 23 new subcategories, such as Orange CARES, Vaccine Outreach, and Testing, to group these pandemic-related requests under distinct subcategories within the *Public Safety* main category. Our analysis indicates the success of the presented model in promptly detecting emerging needs which can notify service providers about the emergence of new categories before the number of anomalies increases extensively.

Moreover, through periodic retraining of the model, the tool gains the ability to recognize whether the system has introduced new categories. This ensures that incoming calls are appropriately classified into recently established categories, rather than defaulting on the generic label of "Other." This adaptability is crucial for staying responsive to evolving circumstances and augmenting the system's precision in categorizing service requests.

To highlight the advantages of periodically retraining the model, Figure 6 illustrates the variation in the number of service calls classified under the category "Other", comparing the data trained before and after the onset of COVID-19, depicted in red and blue lines, respectively. Initially, the dataset was trained on records up to March 1, 2020, marked in red, which did not include specific categories for pandemic-related issues. Subsequently, the model was retrained with data up to April 1, 2021, during which time, new categories to capture COVID-19 related inquiries were established.

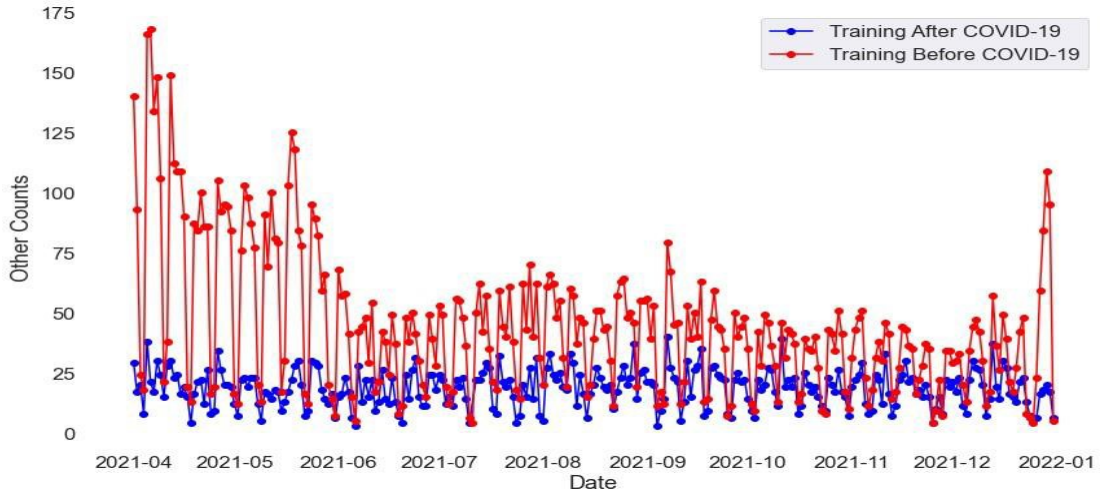


Figure 6. Number of requests classified under "Other" category with training before and after COVID-19

The visualization illustrates a high number of calls classified as "Other" with training before COVID-19, as indicated by the red line, due to the model interpreting all pandemic-related inquiries as anomalies. In contrast, the blue line shows much smaller calls in "Other" category and a consistent pattern after the model is retrained with data up to April 2021, encompassing the first year of the pandemic. The blue line exhibits a small and stable call volume, since COVID-related inquiries are now accurately classified into *Public Safety* main category. Notably, the remaining calls in the "Other" category suggest anomalies stemming from other factors that the model struggles to categorize.

This analysis demonstrates that periodic retraining enables the model to reassign requests previously labeled as "Other" to newly established categories if these inquiries become regular. Hence, the categorization of requests as "Other" is continually refined with each retraining cycle, ensuring it only reflects current anomalies. Thus, the model acts as a valuable tool for service operators, signaling the need to periodically review anomalies and consider creating new categories to effectively classify the evolving types of service calls.

DISCUSSION AND CONCLUSIONS

This work-in-progress paper emphasizes the significant potential of automatic classification of service requests and prompt detection of anomalies to identify emerging needs, especially during disasters. This capability is instrumental in supporting crisis management efforts by enabling public service providers to take timely actions during emergencies. The study's scientific contribution lies in bridging the gap identified in existing literature regarding machine learning approaches for disaster management efforts in service systems. Specifically, while previous studies have primarily focused on classifying service requests into preexisting categories, our proposed model addresses this limitation by incorporating timely anomaly detection.

Our preliminary investigations highlighted a critical issue—the delay in opening new categories for emergent types of service calls. This challenge was especially notable during the COVID-19 period, where it took an extended two months to introduce a category dedicated to COVID-19. Our tool markedly enhances this process, facilitating swift and precise categorization of requests and the identification of irregular requests. Notably, the application of the SVM model demonstrated a 92.19% accuracy rate in classifying service categories. The utility of the approach extends beyond reducing manual classification time for operators; it expedites the recognition and proper categorization of new types of categories. This streamlined approach enhances the overall process by not only automating classification of regular requests but also freeing up human agents to concentrate on detected irregular requests that necessitates a manual evaluation.

The contribution of the study extends beyond merely detecting anomalies related to specific crises; it offers a flexible and adaptable approach applicable to any service system facing disruptions due to evolving service needs during emergencies. Consequently, the presented tool is adaptable and responsive, capable of recognizing and managing newly emerged types of service requests. This underscores the model's extensive applicability in service request management, showcasing its potential to enhance the adaptability and responsiveness of public service systems, particularly in addressing evolving societal needs. This flexibility emphasizes the practicality of the tool in service call systems.

Although the current model excels in classifying main categories, this study faces limitations in classifying requests into subcategories due to the limited number of observations for numerous subcategories. Future research should prioritize improving the model's ability to classify subcategories, which has significant potential in achieving a more nuanced and precise service call categorization and identification of anomalies.

Further improvements could include exploring additional classification models to benchmark against our current method, determining if enhancements are necessary. Moreover, integrating automated alert systems could dynamically notify relevant authorities or response teams upon detection of new types of service needs. Such proactive approaches would ensure swift and targeted responses to emerging challenges, significantly contributing to a more effective overall disaster management strategy.

In conclusion, this study offers a substantial foundation for improving service call categorization systems, with potential widespread implications for emergency response and disaster management. Nonetheless, c ontinuous advancements are essential to fully realize the tool's capabilities and to address the ever-evolving landscape of societal needs. As part of our future work, we plan to incorporate additional classification algorithms, explore different sets of hyperparameters, and implement cross-validation techniques. These efforts will ensure the robustness and generalizability of our findings, enabling a more comprehensive comparison of different modeling approaches and their impact on the classification task.

ACKNOWLEDGMENTS

This research was supported by the University of Vermont's Sustainable Development Policy, Economics, and Governance (SDPEG) interdisciplinary Ph.D. program, and by a grant from the National Science Foundation (NSF-SCC #1952792). The authors would like to thank Orange County, Florida 311 for their collaboration and for sharing their datasets. We also acknowledge the primary and equal contribution of the first and second authors, Hakan Unveren and Ayesh Abu Lehyeh, to this work.

REFERENCES

Aipe, A., Mukuntha, N., Ekbal, A., & Kurohashi, S. (2018). Deep learning approach towards multi-label classification of crisis related tweets. In Proceedings of the 15th ISCRAM Conference.

Alam, F., Ofli, F., & Imran, M. (2019). CrisisDPS: Crisis Data Processing Services. In Proceedings of the 16th ISCRAM Conference.

Apidopoulos, V., Poggio, T., Rosasco, L., & Villa, S. (2022). Iterative regularization in classification via hinge loss diagonal descent. *arXiv preprint arXiv:2212.12675*.

Bianchini, D., De Antonellis, V., Melchiori, M., & Salvi, D. (2006). Semantic-enriched service discovery. 22nd International Conference on Data Engineering Workshops (ICDEW'06).

Brogi, A., Corfini, S., Aldana, J. F., & Navas, I. (2006). Automated discovery of compositions of services described with separate ontologies. International Conference on Service-Oriented Computing. Berlin, Heidelberg: Springer Berlin Heidelberg.

Buntain, C., McCreadie, R., & Soboroff, I. (2022). Incident Streams 2021 Off the Deep End: Deeper Annotations and Evaluations in Twitter. In Proceedings of the 19th ISCRAM Conference.

Cheng, S., Ganapati, S., Narasimhan, G., & Yusuf, F. B. (2022). A machine learning-based analysis of 311 requests in the Miami-Dade County. *Growth and Change*, *53*(4), 1627-1645.

City of Chicago. (2018). *Mayor Emanuel Launches Modernized Chicago 311 System*. Press Room. https://www.chicago.gov/city/en/depts/mayor/press_room/press_releases/2018/december/Chicago311.html

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*, 273-297.

Cutway, A. (2020). Timeline: The Spread of coronavirus in Florida. *Click Orlando.com*. https://www.clickorlando.com/news/local/2020/03/20/timeline-the-spread-of-coronavirus-in-florida/

Filip, R., Gheorghita Puscaselu, R., Anchidin-Norocel, L., Dimian, M., & Savage, W. K. (2022). Global challenges to public health care systems during the COVID-19 pandemic: a review of pandemic measures and problems. *Journal of Personalized Medicine*, *12*(8), 1295.

Galar, M., Fernández, A., Barrenechea, E., Bustince, H., & Herrera, F. (2011). An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recognition*, *44*(8), 1761-1776.

Gualtieri, J. A., & Cromp, R. F. (1999). Support vector machines for hyperspectral remote sensing classification. In 27th AIPR workshop: Advances in computer-assisted recognition (Vol. 3584, pp. 221-232). SPIE.

Hashemi, M. (2022). Automatic Type Detection of 311 Service Requests Based on Customer Provided Descriptions. *Applied Artificial Intelligence*, *36*(1), 2073717.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.

Hsu, C.-W., & Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, *13*(2), 415-425.

Kersten, J., Bongard, J., & Klan, F. (2021). Combining supervised and unsupervised learning to detect and semantically aggregate crisis-related twitter content. In Proceedings of the 18th ISCRAM Conference.

Kontokosta, C., Hong, B., & Korsberg, K. (2017). Equity in 311 reporting: Understanding socio-spatial differentials in the propensity to complain. *arXiv preprint arXiv:1710.02452*.

Li, H., Caragea, D., & Caragea, C. (2021). Combining self-training with deep learning for disaster tweet classification. In Proceedings of the 18th ISCRAM Conference.

Li, H., Li, X., Caragea, D., & Caragea, C. (2018). Comparison of word embeddings and sentence encodings as generalized representations for crisis tweet classification tasks. In Proceedings of ISCRAM Asia Pacific 2018 Conference.

O'Brien, D. T. (2016). 311 hotlines, territoriality, and the collaborative maintenance of the urban commons: Examining the intersection of a coproduction policy and evolved human behavior. *Evolutionary Behavioral Sciences*, *10*(2), 123.

Orange County Government, Florida. (n.d.). Retrieved 02/02/2024 from https://www.ocfl.net/AboutUs.aspx

Pamukcu, D., Zobel, C., & Ge, Y. (2022). A Data Envelopment Analysis-based approach for managing performance of public service systems during a disaster. In Proceedings of the 19th ISCRAM Conference.

Pamukcu, D., Zobel, C. W., & Ge, Y. (2021). Analysis of Orange County, Florida 311 system service requests during the COVID-19 pandemic. In Proceedings of the 18th ISCRAM Conference.

Samuel, L. (2019). *Cities Embrace New and Improved 311 Services*. Data-Smart City Solutions Retrieved 22/01/2024 from https://datasmart.hks.harvard.edu/news/article/cities-embrace-new-improved-311-services

Schwester, R. W., Carrizales, T., & Holzer, M. (2009). An examination of the municipal 311 system. *International Journal of Organization Theory & Behavior*, *12*(2), 218-236.

Shah, S. A., Seker, D. Z., Hameed, S., & Draheim, D. (2019). The rising role of big data analytics and IoT in disaster management: recent advances, taxonomy and prospects. *IEEE Access*, *7*, 54595-54614.

Shalev-Shwartz, S., Singer, Y., & Srebro, N. (2007). Pegasos: Primal estimated sub-gradient solver for svm. In Proceedings of the 24th International Conference on Machine Learning.

Wang, L., Qian, C., Kats, P., Kontokosta, C., & Sobolevsky, S. (2017). Structure of 311 service requests as a signature of urban location. *PloS one*, *12*(10), e0186314.

Wiegmann, M., Kersten, J., Klan, F., Potthast, M., & Stein, B. (2020). Analysis of detection models for disaster-related tweets. *Analysis of Detection Models for Disaster-Related Tweets*, 872-880.

Winbourne, J. (2022, October 18, 2022). *311 Services — A key Step in the Global Transition to e-Government Services* https://www.w-llc.com/311-services-a-key-step-in-the-global-transition-to-e-government-services/

Wu, W. N. (2020). Determinants of citizen-generated data in a smart city: Analysis of 311 system user behavior. *Sustainable Cities and Society*, *59*, 102167.

Scenario planning to enable foresight in crisis management

Marte Luesink
Nederlands Instituut Publieke Veiligheid/
Leiden University
marte.luesink@nipv.nl

Jeroen Wolbers
Leiden University
j.j.wolbers@fgga.leidenuniv.nl

Menno van Duin
Nederlands Instituut Publieke Veiligheid
menno.vanduin@nipv.nl

Sanneke Kuipers
Leiden University
s.l.kuipers@fgga.leidenuniv.nl

ABSTRACT

In today's interconnected society, we are witnessing crises traversing into new domains and effortlessly cross boundaries. Therefore, it is crucial to develop adequate foresight in order to chart a course of action. In this paper, we propose *scenario planning* as a key technique. Based on an explorative study of 30 interviews with representatives from all Dutch safety regions, we answer the question: *in what ways can scenario planning facilitate foresight in crisis information management?* Our results indicate that all regions in the Netherlands are investing in training and capacity building for scenario planning, but there is no consensus on the chosen methods and institutionalization. Qualitative approaches to scenario planning are dominant, while opportunities arise for quantitative approaches. We propose *real-time* and *strategic foresight* as a promising research agenda, and suggest ways for information systems research to develop data-driven tools that may help chart the impact of future contingencies.

Keywords

Scenario planning, foresight, polycrisis, sensemaking, information systems.

INTRODUCTION

When a crisis emerges, agencies are confronted with an acute threat that requires immediate action. Time pressure, incomplete information, and the demand for prompt decision-making add to the sense of urgency crisis managers experience (Rosenthal & Kouzmin, 1997). This urgency pulls crisis managers into the 'here and now', which often results in a reactive modus operandi with a limited time horizon (Jaques, 2010). Consequently, persons managing a crisis often find it difficult to look ahead and incorporate future contingencies in their decision-making. This short-term focus is also reflected in the academic debate, which tends to focus on the response phase. Crisis research is packed with single case studies that offer in-depth analyses of the response to a specific crisis (Wolbers et al., 2021). Yet, the dominant focus on single case studies limits our view on multiple dimensions or spill-over effects of crises, which could be problematic. Like the saying has it 'trouble always comes in threes'.

In late 2022, the Financial Times columnist and historian Adam Tooze coined the concept of a 'polycrisis'. Tooze argued that the world is enmeshed in a variety of crises that are entangled and strengthen each other. He described the interplay between the COVID-19 pandemic, the war in Ukraine and the energy, cost-of-living and climate crises (Tooze, 2022). In a polycrisis, different crises interact so that the whole is even more overwhelming than the sum of its parts (c.f. Jervis, 1998). While the term polycrisis resonates predominantly in the public realm, in the academic debate we are familiar with typologies that describe how crises become too complex to manage when crossing geographical or functional boundaries. We speak of a transboundary crisis *"when the functioning of multiple, life-sustaining systems is acutely threatened"* (Boin & Rhinard, 2008, p. 3).

The manifestation of a transboundary crisis tends to morph as its impact prolongs. When a transboundary crisis impacts multiple domains, the nature of the crisis itself may also change. For example, the alleged Russian 'Not-Petya' cyber-attack during the Russia-Ukrainian war showed how the crisis almost instantly traversed domains around the world. Within minutes the bug rendered both the FedEx distribution system and Maersk worldwide

container shipping network unusable (Greenberg, 2018). Such escalation pathways point us to the necessity for crisis managers to look ahead for ways to address a crisis traversing into new systems and domains. They require crisis managers to develop *strategic foresight*, the capacity of decision-makers to chart the organization's future course of action (Vecchiato, 2012). Developing strategic foresight entails identifying factors that are likely to induce future changes, while also dealing with those changes by deciding on currently appropriate organizational responses (Iden et al., 2017).

Developing strategic foresight during a crisis is not easy. In an extreme situation, limited awareness of possible escalation pathways may result in the 'collapse of the future', when the long-term perspective disappears from crisis managers' time horizon (Pot et al., 2022). Short-term decisions can have consequences that stretch far beyond the critical period. If crisis managers do not take the long-term into account, they will miss cues that point to opportunities for de-escalation, or miss cues that point to the escalation of the crisis into other domains. Therefore, crisis managers need to create moments of reflection and look for decision-support tools that can help them refocus on the long-term in order to foster strategic foresight. Foresight support systems may play a crucial role here as they feature a combination of information system and decision-support tools (Von der Gracht et al., 2015).

In this paper, we highlight the role that scenarios can play to embrace uncertainty and focus on the long-term (Wilkinson & Kupers, 2013; Wright, 2005). Scenarios are a tool to help decision-makers chart possible future developments by *"making the unexpected expectable by developing a set of narratives that incorporate uncertainty"* (Wright, 2005, p. 95). In developing alternative futures, crisis managers can challenge dominant frames in the current crisis. As such, scenarios form prospective sensemaking devices that can facilitate structuring the future in order to identify opportunities for action (MacKay, 2009; Burt & van der Heijden, 2003).

Given that scenario planning is a trend in crisis management, it is helpful to explore and chart the ways in which scenarios are used. We relate the relevant literature on scenario planning and use that to map the current state-of-the-art developments in scenario planning in the Netherlands. We pose the question: *in what ways can scenario planning facilitate foresight in crisis information management?* We base our analysis on an explorative study consisting of 30 interviews with regional crisis agencies in the Netherlands. Based on this analysis, we explore the goals, different scenario methods and design choices that characterize scenario planning in said agencies. This enables us to reflect on state-of-the-art developments and identify an agenda for future research that may inspire crisis scholars investigate ways in which to support scenario planning and the development of foresight. For information systems scholars, we identify opportunities to complement qualitative scenario processes by integrating quantitative methods to support real-time and strategic foresight.

DEVELOPING SCENARIOS TO ENABLE FORESIGHT

From the very first impact of the crisis onwards, it can be helpful for crisis managers to develop strategic foresight to foresee and plan for possible escalation pathways and chart a future course of action (Vecchiato, 2012). In order to rethink plausible developments and generate options for action, many crisis teams are currently investing in scenario methods (Ramírez & Wilkinson, 2014). Scenarios are a helpful tool for ordering one's perceptions about alternative futures to inform decision-making (Schwartz, 1996). They can function as a tool to enable prospective sensemaking (Wright, 2005), which is referred to as the process in which actors anticipate future developments and construct expectations about the future (Sandberg & Tsoukas, 2015). This is similar when crisis managers engage in the process of scenario development, where they engage in continuous inquiry and re-evaluation of interpretations that helps them to make sense of a rapidly changing context (Ansell & Boin, 2019).

Different concepts are used to describe the process of generating scenarios, such as scenario planning, scenario thinking and scenario development. In this paper we use the definition provided by Bishop et al. (2007) who suggest that scenario planning is part of a foresight study, where scenarios help to *"create actual stories about the future. Scenario planning is a far more comprehensive activity, of which scenario development is one aspect"* (Bishop et al., 2007, p. 6). As scenario planning is booming in business continuity and crisis management, a multitude of methods are being developed, contributing to a wide array of concepts and approaches (Martelli, 2001). Surprisingly, in the crisis management literature there is scarce attention for scenario planning and foresight.

Scenario methods can be categorized into different schools representing either *plausibility* or *probability* based approaches (Bradfield et al., 2005). *Plausibility-based approaches* are qualitative approaches that aim to explore possible futures and focus on facilitating prospective sensemaking. In this approach, scenarios are used as prospective sensemaking devices that help to transition from a retrospective to a future-oriented focus by patterning cues and plots into frames to imagine alternative futures (Wright, 2005; Ramírez et al., 2013; Wolbers & Boersma, 2013).

A well-known approach is the 'intuitive logics' approach, first developed by Shell in the 60s to envision trends in the energy sector beyond the year 2000. In contrast, *probability-based approaches* are quantitative approaches that aim to predict the future by modelling historical data to extrapolate trends and determine what consequences the probability of unprecedented events has on these trends (Bradfield et al., 2005). Approaches like 'Trend-Impact Analysis' and 'Cross-Impact Analysis' are used to determine conditional or proportional probabilities to calculate impact and generate a range of alternative futures. During Covid-19 we have seen the extensive use of quantitative methods and modelling, primarily regarding infection rates, covid related deaths and ICU-capacity, to support strategic policy decisions (i.e. Shinde et al., 2020). Such data can help foster grounds for decision making, yet evaluations stressed the importance of the validity of data and the need to keep assessing whether assumptions are still valid as the crisis progresses (Dutch Safety Board, 2023).

What becomes clear from these different approaches of scenario planning is that they serve different goals and outcomes (Burt & Van der Heijden, 2003). For the use of scenarios in crisis management, it matters whether scenarios are used in the acute phase of a crisis to enhance option awareness about escalation pathways (Liu et al., 2011; Pfaff et al., 2023), or if scenarios are used to devise alternative futures in a slower paced or creeping crisis (Sætren et al., 2023, Mascio et al., 2020). In the acute phase, when there is little time, intuition will play a key role in scenario developers' ability to recognize and analyse cues and patterns (Barlett et al., 2013). In a slower or creeping crisis, a more structured form of sensemaking can be based on the extrapolation of risk calculations from databases. Information systems support can be helpful to create these databases. Glenn (2015), for example, argues the use of collective information systems as a way of using and supporting foresight methods, and Keller et al. (2015) propose a foresight database to serve as a 'knowledge pool' for foresight support systems. Needless to say, the quality of such risk calculations will depend on the availability of relevant (historical) data.

METHODS

In order to map the current developments in scenario planning amongst crisis actors in the Netherlands, the primary author conducted an explorative study into the use of scenario planning in all 25 safety regions in the Netherlands (Luesink & Boin, 2023). A safety region is a regional crisis management organization facilitating interdisciplinary cooperation of emergency services during a (major) incident or crisis. For the study, 30 respondents were interviewed, who were selected based on their involvement with scenario planning in education, training and practicing within the safety regions. This resulted in 24 interviews with representatives from the 25 safety regions (two regions share their staff). The study is thus representative of crisis management at the regional level in the Netherlands. In addition, six interviews were held with relevant crisis management partners who also teach scenario planning to crisis managers: the Police Academy, National Academy for Crisis Management, the Public Health Academy, Intelligence & Security Academy of the Ministry of Defense, a consultancy firm, and the Netherlands Institute for Public Safety. In semi-structured, open interviews, respondents were asked to reflect on the process of implementing scenario planning trainings between 2016 and 2022. They reflected on the choices made in this process and shared their experiences with scenario planning in crisis situations. The interviews lasted around 60 minutes. All interviews were recorded and transcribed. The transcripts were coded in the manner of grounded theory (Charmaz, 2014). During the coding process, the authors discussed the interpretation of the data and possible categories with each other on a weekly basis in the course of a month. The categories were identified that included the characteristics of the scenario processes, the goals that were set for scenario planning, the design and methods used, and the intended audience for the scenarios (Luesink & Boin, 2023). Together this created a solid overview of the state-of-the-art developments in scenario planning in the Netherlands.

FINDINGS

Goal & Purpose

Scenario planning gained traction in the Netherlands in the wake of the Covid-19 pandemic, as crisis teams began to chart possible futures to gain a sense of control over the prolonged crisis. Respondents point to a desire to 'get out in front of the crisis' in order to organize a proactive, rather than a reactive, response (Luesink & Boin, 2023). While a multitude of regions had already invested in teams that structurally monitored safety related developments in the 'cold' phase, the pandemic gave the quest for foresight a final push. Across all safety regions, agencies rapidly allocated human resources, hired trainers and established teams to carry out scenario planning in order to advise decision-makers. As of now, most regions and national crisis centers in the Netherlands use some form of scenario planning in crisis management. This rapid scaling up has given rise to a great diversity of applied scenario methodologies.

Overall, respondents noted two main purposes of scenario development. A primary goal is to create awareness about possible future developments. Scenario teams aimed to create a shared idea of possible futures so they could provide crisis managers with plausible grounds on which to devise a future course of action in an uncertain context. Some respondents focused on getting a grip on an uncertain situation by stating that *"the aim is to keep some level of control over the situation and to recognize how a crisis is developing."* Others emphasized the need *"to walk through a possible scenario together to make sure we are all on the same page"* (Luesink & Boin, 2023, translation ours).

A secondary, complementary goal that respondents identified is anticipating possible future developments and (proactively) taking measures accordingly. This goal corresponds with the need to formulate appropriate organizational responses (Iden et al., 2017). Even though the situation might be highly uncertain, and little verified information is available, decisions and measures can be taken in consideration of plausible escalation pathways of the situation at hand. A respondent stated that the purpose of scenario planning is *"to recognize the scenario we are in and to identify preventive and reactive measures"* (Luesink & Boin, 2023). The first goal can be described with the 'what if' question. The second goal entails the subsequent 'then what' question.

These two main purposes of scenario development, as identified by respondents, correspond with the dual tasks related to fostering strategic foresight (Iden et al., 2017): anticipation and taking (preventive) measures. The first step is exploring the situation at hand and identifying factors that could influence future developments. The second step is imagining and, possibly, preparing intervention measures to facilitate desirable and prevent undesirable developments.

Design & Methodology of Scenario Development

In the literature on scenario planning, several authors including Martelli (2001) decry the 'methodological chaos' in the field that stems from a plethora of scenario methods. Luesink & Boin (2023) find a similar diversity of scenario methods in the Netherlands. This is not necessarily a problem. After all, there seems to be no 'one size fits all' scenario method, since the selection of the most appropriate method should be guided by goal and organizational context of the scenario developers (Burt & Van der Heijden, 2003).

We provide an overview of the different scenario methods used in the crisis management landscape in the Netherlands in Table 1. These are four main methods taught in safety regions during scenario development training: 'Forward Thinking', 'IRS model', 'Imagination', and the 'Shell method' (i.e. intuitive logics) (see Table 1 for an explanation of each). These four scenario methods are all qualitative methods yet differ from one another in several ways. An important difference is that the 'Forward Thinking', 'IRS model' and 'Imagination' methods were designed or adapted to cater to the needs of crisis managers, whereas the 'Shell method' was designed to facilitate long term strategic thinking in a competitive business environment. The scenario formulation process of the first three methods also takes up less time than the 'Shell method', thereby offering a good fit for more high-paced crisis situations.

The 'Forward Thinking' and the 'Shell method' are both focused on strategic crisis processes. By engaging in these methods of scenario planning, crisis managers can get a better handle on the inherent uncertainty of a crisis situation and open their minds to multiple possible futures. The 'IRS model' and 'Imagination method', on the other hand, are characterized by an action-oriented purpose whereby the scenarios are intended to facilitate the identification of options for intervention and possible measures. Particularly in the 'IRS model', the scenarios are not the main outcome. Rather, scenarios aid in imagining a possible future state in order to formulate possible measures for intervention during the crisis.

These scenario methods portray a different operationalization of the dual tasks of strategic foresight. Whereas all methods engage with the situation at hand to explore possible futures, and subsequently identify options for response, the way scenarios help to identify these options differs. The 'Forward thinking', 'IRS model' and 'Imagination' methods demarcate the future through a best-case and worst-case scenario: developments that fall "between" these two scenarios are deemed most likely and should be considered when making decisions. In the 'Shell method', scenarios are created by ranking factors on uncertainty and their impact. Consequently, the scenarios are created based on the perceived uncertainties of the situation at hand.

Table 1. Overview of four common scenario methods as used by safety regions (between 2016 and 2022) to train crisis managers in scenario development.

Method	Source/based on	Crisis management level	Purpose	Approach	Process	Outcome
Forward thinking	Crisis management training consultancy	Strategic level	Coping with uncertainty	Strategic analysis	1. Analyze the current situation 2. Develop a favorable and unfavorable scenario 3. Identify strategic dilemmas for each scenario 4. Prioritize strategic decisions	Two scenarios and prioritized strategic decisions
IRS model	National Police	Operational and tactical level	Identifying and prioritizing measures	Risk analysis	1. Identify risks 2. Rank risks based on probability and impact 3. Develop a best, realistic and worst-case scenario for the high priority risks 4. Propose measures (preventive and reactive) for one or more scenarios	Risk matrix
Imagination	National Academy of Crisis Management	Operational, tactical, and strategic level	Identifying opportunities for intervention	Exploration of different future pathways	1. Analyze the current situation 2. Develop best, realistic and worst-case scenarios 3. Identify scenario drivers 4. Describe opportunities for intervention	Three scenarios, scenario drivers and opportunities for intervention
Shell method	Schwartz (1996)	Strategic level	Challenging mindset of decision makers	Exploration of possible alternative futures	1. Identify the focal issue or decision 2. Identify key forces in the micro-environment 3. Identify driving forces in the macro-environment 4. Rank the key and driving forces based on uncertainty and importance 5. Place the two key uncertainties on the axes of the 2 x 2 matrix 6. Write the four scenarios 7. Explore the implications of each scenario 8. Select indicators and triggers for monitoring	Four scenarios and robust response options

Data retrieved from interviews with respondents from the scenario methods; 1. Nijhof Teamadvies (consultancy), 2. National Police, 3. National Academy of Crisis Management, and 4. the Shell method as described by Schwartz (1996).

Organization and Institutionalization

A key difference noted in the safety regions in the Netherlands concerned the organizational level at which scenario development should take place and, thus, which crisis managers should be trained. In safety regions, crisis management is organized at three levels: the operational level (focused at the crisis response on the incident ground), the tactical level (focused at supporting the operational level and/or advising the strategic level) and the strategic level (focused at longer-term decisions by public local/regional authorities). Twenty-two respondents regarded the tactical level, and particularly the leader of that team, as the most suitable context for the development of scenarios. In the words of one respondent: *"the crisis team at the tactical level is most suited for scenario planning because they have more time to develop scenarios than the operational level and because the team takes both the operational and strategic level into account"* (Luesink & Boin, 2023). The crisis managers at the tactical level are expected to consider information from the operational level and translate it to long term effects and strategic dilemmas. The subsequent scenarios can be focused on either the operations or the strategy of the crisis response.

Many respondents pointed out challenges for scenario planning at the operational and strategic levels. Respondents argued that the operational level already has a great deal of tasks to complete in very little time, making scenario planning too much to ask even if it could be to their own advantage. A concern regarding the strategic level that several respondents shared is that the strategic crisis team would focus too much on the operational level instead of sticking to strategic dilemmas and decisions. Indeed, previous studies have often established that strategic teams which engage with operational units frequently fall prey to 'operational fixation' (Boin et al., 2017; Wolbers, 2022). Moreover, the majority view among respondents is that tactical crisis managers can strengthen their advice to strategic decision makers through the use of scenarios. *"The strategic level has advisors at the tactical level which support them in their decision making through scenario planning"*, in the words of one respondent (Luesink & Boin, 2023).

The discussion of who should engage in scenario planning concerns not only the level at which it should be done. We also noted different approaches to where in the organization to embed scenario planning: integrated in the work of every crisis manager of the responsibility of a designated scenario team? A quarter of the respondents argue in favor of keeping scenario development closely related to other crisis management processes such as sensemaking and decision-making, and thus integrated within the crisis team. Sixteen respondents, see advantages to having a separate scenario team. *"A separate scenario team would have more time to develop scenarios during a crisis and could consist of crisis managers who have more affinity with foresight"*, as one respondent stated (Luesink & Boin, 2023). Eight of the sixteen respondents, however, argue that a scenario team would only work during complex and/or protracted crises. The utility of a scenario team during shorter, less complex crises is drawn into question by these respondents, who note that by the time scenarios are developed and relayed, acute decisions will have already been taken and the crisis might already be under control. Setting up rapid scenario teams could also make the number of actors involved in the crisis response disproportionately high, resulting in higher coordination costs (Luesink & Boin, 2023).

CONCLUSION AND THE ROAD AHEAD: DEVELOPING REAL-TIME AND STRATEGIC FORESIGHT

After the COVID-19 pandemic, scenario planning and foresight have received much attention. Actors across safety regions and national crisis agencies in the Netherlands have invested in training and capacity building. Organizations are fully engaged with anchoring scenario methods in crisis response processes and structures through scenario teams. Currently, however, the variation in opinions and approaches between the regions and organizations is telling. No consensus between the crisis organizations exists on the questions of when, how or by whom scenario planning should be done. Flexibility and adaptability are key in crisis management, so scenario practices do not have to be pinned down completely – still, respondents attest to missing common ground. Within, but especially between, the 25 safety regions, sharing a common understanding of the goal of scenario planning and methods is deemed most applicable for different types of scenario development.

Looking towards the future ourselves, we note two pathways of interest that will help in advancing scenario planning: *real-time foresight* and *strategic foresight*.

Real-Time Foresight

Scenario planning is often related to broader processes of risk analysis and monitoring. Real-time foresight concerns the process by which crisis organizations *"concurrently prepare for and respond to a crisis"* (Wells et al., 2021, p. 35). It entails foreseeing future developments that can be acted upon by crisis managers. In that sense, real-time foresight links to the third level of situational awareness, which entails the anticipation of future developments, based on a detailed understanding of the situation at hand (Endsley, 1995).

To facilitate the process of real-time foresight, we can look to foresight support systems, such as dashboards, collective intelligence systems, data-mining, and web 2.0 approaches (Von der Gracht et al., 2015). When organizations are able to access or construct databases that help to identify trends and prognoses, it could deliver indicators that can be monitored. We have already seen examples in the Netherlands of dashboards monitoring the influx of Ukrainian refugees and virtual assistants in crisis teams to provide actionable advice.

The difficulty in monitoring is knowing what to look for. No crisis unfolds exactly like its predecessors. The explorative study of scenario planning demonstrates the current focus on qualitative, intuitive logics scenario planning in crisis management (Luesink & Boin, 2023). This type of scenario planning can be useful for analyzing the current situation, exploring possible future developments and identifying indicators which can be monitored. Once relevant indicators have been established, quality and availability of data become the next obstacles. Very few databases will likely be available to conduct trend analysis at the start of a crisis. However, a prolonged crisis, like climate change or migration, can give the necessary time to construct relevant databases.

Strategic Foresight

Developing strategic foresight is the major challenge of the era of the polycrisis. While most scenario teams are still focused on a singular crisis, recent developments require that attention to the possibility that crises may entangle and strengthen each other. The polycrisis points to the importance of foreseeing escalation pathways and imagining the relationship between different domains. For example, the World Economic Forum Global Risk Perception Survey (2023) has identified a range of escalation pathways between crises. Looking ahead for ways to address a crisis that might traverse into new systems and domains can be facilitated by using specific scenario methods that help to foster strategic foresight (Vecchiato, 2012). To find the right approach, researchers and practitioners need to explore a variety of scenario tools that may help 'zoom out' from the acute threats of a single crisis and let crisis managers consider the interrelated nature of our worldwide networks and infrastructures. As systems become ever more connected, and the destructive potential and cascading effects of crises increase accordingly, scenario planning may play a crucial role in supporting the anticipation and mitigation of future crises.

The Road Ahead: Developing Information Systems for Scenario Planning

Scholars widely recognize the unique capabilities of technology to enhance foresight (Marinković et al., 2022). Information systems research provides promising avenues into data-driven and quantitative tools that help to assess the relevance and possible impact of future contingencies (Von der Gracht et al., 2015). As we witness the increasing availability of new data sources, such as social media, web 2.0 and open data, in combination with machine learning algorithms and computing capacity, a rich set of possibilities is opening up to advance the capacity for foresight (Geurts et al., 2021). It is time to take the next steps to develop systems that provide collective intelligence capabilities in order to increase the analytical leverage for crisis managers struggling to engage in foresight processes and scenario planning. Especially the integration of qualitative and quantitative methods can strengthen foresight processes. Examples of such steps in different contexts can be seen in, for instance, the Regional Innovation System (Keller et al., 2015) and the Risk Assessment and Horizon Scanning project (Durst et al., 2015). Through such information systems foresight processes are enhanced, by offering consistency checks, collaboration of geographically dispersed experts, and database construction.

Still, we are not there yet. First, academic research on foresight seems to be primarily focused on business strategy, business continuity and innovation, whereas research on foresight in crisis management is scarce. Second, a persistent challenge is to explicate the required capabilities of participants in foresight processes (Durst et al., 2015). Without a clearer idea on what foresight processes entail in crisis management, what kind of information is required, and what kind of assessments help to foresee escalation pathways, it remains hard to offer a fitting technological infrastructure. Our research indicates that training in scenario planning is currently primarily focused on qualitative methods, while adequate risk assessment and monitoring require the integration of quantitative tools and databases. Such integration may help to validate and test the intuitive future contingencies that crisis managers foresee. Furthermore, it remains a key challenge to assemble the right information in the middle of a crisis. It would be interesting to explore whether other countries face similar challenges or have already found solutions. A lot of potential is offered by unlocking new data sources, such as artificial intelligence, web 2.0 and open data, in order to provide a richer and more responsive data environment to the crisis manager. To enable both real-time and strategic foresight in the midst of a crisis, we need to combine best of both worlds to design a fitting technological infrastructure. We hope the ISCRAM community can play a leading role in fostering such innovative capacity for crisis managers struggling to set a future course of action in the rugged and unpredictable landscape of the polycrisis.

REFERENCES

Ansell, C., & Boin, A. (2019). Taming Deep Uncertainty: The Potential of Pragmatist Principles for Understanding and Improving Strategic Crisis Management. *Administration & Society, 51*(7), 1079-1112.

Barlett, K., Nolan, M., & Marraffino, A. (2013). Intuitive Sensemaking: From Theory to Simulation Based Training. In D. D. Schmorrow, & C. M. Fidopiastis (Eds.), *Foundations of Augmented Cognition* (pp. 3-10). Springer.

Bishop, P., Hines, A., & Collins, T. (2007). The current state of scenario development: an overview of techniques. *Foresight, 9*(1), 5-25.

Boin, A., & Rhinard, M. (2008). Managing transboundary crises: What role for the European Union? *International Studies Review, 10*(1), 1-26.

Boin, A., Stern, E., & Sundelius, B. (2017). *The politics of crisis management: Public leadership under pressure*. Cambridge University Press.

Bradfield, R., Wright, G., Burt, G., Cairns, G., & van der Heijden, K. (2005). The origins and evolution of scenario techniques in long range business planning. *Futures, 37*(8), 795-812.

Burt, G., & van der Heijden, K. (2003). First steps: Towards purposeful activities in scenario thinking and future studies. *Futures, 35*(10), 1011-1026.

Charmaz, K. (2014). *Constructing Grounded Theory* (2nd ed.). SAGE.

Dutch Safety Board. (2023). *Aanpak coronacrisis. Deel 3: januari 2020 – september 2022*. https://onderzoeksraad.nl/onderzoek/aanpak-coronacrisis-deel-3/

Durst, C., Durst, M., Kolonko, T., Neef, A., & Greif, F. (2015). A holistic approach to strategic foresight: A foresight support system for the German Federal Armed Forces. *Technological Forecasting & Social Change, 97*, 91-104.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors, 37*(1), 32-64.

Geurts, A., Gutknecht, R., Warnke, P., Goetheer, A., Schirrmeister, E., Bakker, B., & Meissner, S. (2021). New perspectives for data-supported foresight: The hybrid AI-expert approach. *Futures & Foresight Science, 4*(1), 1-14.

Glenn, J.C. (2015). Collective intelligence systems and an application by The Millenium Project for the Egyptian Academy of Scientific Research and Technology. *Technological Forecasting & Social Change, 97*, 7-14.

Greenberg, A. (2018, August 22). The untold story of NotPetya, the most devastating cyberattack in history. *Wired*.

Iden, J., Methlie, L. B., & Christensen, G. E. (2017). The nature of strategic foresight research: A systematic literature review. *Technological Forecasting & Social Change, 116*, 87-97.

Jaques, T. (2010). Reshaping crisis management: The challenge for organizational design. *Organization Development Journal*, *28*(1), 9.

Jervis, R. (1998). *System effects: Complexity in political and social life*. Princeton University Press.

Keller, J., Markmann, C., & Von der Gracht, H. A. (2015). Foresight support systems to facilitate regional innovations: A conceptualization case for a German logistics cluster. *Technological Forecasting & Social Change, 97*, 15-28.

Liu, Y., Moon, S. P., Pfaff, M. S., Drury, J. L., & Klein, G. L. (2011, May). Collaborative option awareness for emergency response decision making. *Proceedings of the 8th International ISCRAM Conference, 2011, Lisbon, Portugal.*

Luesink, M., & Boin, C. (2023). *Trainen in scenariodenken binnen veiligheidsregio's: Een eerste inventarisatie.* Nederlands Instituut Publieke Veiligheid.

MacKay, R. B. (2009). Strategic foresight: counterfactual and prospective sensemaking in enacted environments. In L. A. Costanzo, & R. B. MacKay (Eds.), *Handbook of Research on Strategy and Foresight* (pp. 90-112). Edgar Elgar Publishing Limited.

Marinković, M., Al-Tabbaa, O., Khan, Z., & Wu, J. (2022). Corporate foresight: A systematic literature review and future research trajectories. *Journal of Business Research, 144*, 289-311.

Martelli, A. (2001). Scenario building and scenario planning: state of the art and prospects of evolution. *Futures*

Research Quarterly, 17, 57-70.

Mascio, F. D., Natalini, A., & Cacciatore, F. (2020). Public administration and creeping crises: Insights from COVID-19 pandemic in Italy. *The American Review of Public Administration, 50*(6-7), 621-627.

Pfaff, M. S., Klein, G. L., Drury, J. L., Moon, S. P., Liu, Y., & Entezari, S. O. (2013). Supporting complex decision making through option awareness. *Journal of Cognitive Engineering and Decision Making, 7*(2), 155-178.

Pot, W., Scherpenisse, J., & 't Hart, P. (2022). Robust governance for the long term and the heat of the moment: Temporal strategies for coping with dual crises. *Public Administration, 101*(1), 221-235.

Ramírez, R., & Wilkinson, A. (2014). Rethinking the 2 x 2 scenario method: Grid or frames? *Technological Forecasting & Social Change, 86*, 254-264.

Ramírez, R., Österman, R., & Grönquist, D. (2013). Scenarios and early warnings as dynamic capabilities to frame managerial attention. *Technological Forecasting and Social Change, 80*(4), 825-838.

Rosenthal, U., & Kouzmin, A. (1997). Crises and Crisis Management: Toward Comprehensive Government Decision Making. *Journal of Public Administration Research and Theory, 2*, 277-304.

Sætren, G. B., Vaag, J. R., Hansen, I. F., & Bjørnfeld, G. A. (2023). Situational awareness in a creeping crisis: How the initial phases of the COVID-19 pandemic were handled from a crisis management perspective in the Nursing Home Agency in Oslo. *Journal of Contingencies and Crisis Management, 31*, 545–559.

Sandberg, J., & Tsoukas, H. (2015). Making sense of the sensemaking perspective: Its constituents, limitations, and opportunities for further development. *Journal of Organizational Behavior, 36*(S1), 6-32.

Schwartz, P. (1996). *The Art of the Long View: Planning for the Future in an Uncertain World*. Crown Business.

Shinde, G. R., Kalamkar, A. B., Mahalle, P. N., Dey, N., Chaki, J., & Hassanien, A. E. (2020). Forecasting models for coronavirus disease (COVID-19): a survey of the state-of-the-art. *SN computer science, 1*, 1-15.

Tooze, A. (2022, October 28). Welcome to the World of Polycrisis. *Financial Times*.

Vecchiato, R. (2012). Environmental uncertainty, foresight and strategic decision making: An integrated study. *Technological Forecasting & Social Change, 79*, 436-447.

Von der Gracht, H. A., Bañuls, V. A., Turoff, M., Skulimowski, A. M., & Gordon, T. J. (2015). Foresight support systems: The future role of ICT for foresight. *Technological Forecasting and Social Change, 97*, 1-6.

Wells, E. M., Cummings, C., Klasa, K., Trump, B. D., Cegan, J., & Linkov, I. (2021). Real-time Anticipatory Response to COVID-19: A Novel Methodological Approach. In I. Linkov, J. M. Keenan, & B. D. Trump (Eds.), *COVID-19: Systemic Risk and Resilience* (pp. 35-59). Springer.

Wilkinson, A., & Kupers, R. (2013). Living in the Futures. *Harvard Business Review, 91*(5), 118-127.

Wolbers, J., & Boersma, K. (2013). The common operational picture as collective sensemaking. *Journal of Contingencies and Crisis management, 21*(4), 186-199.

Wolbers, J., Kuipers, S., & Boin, A. (2021). A systematic review of 20 years of crisis and disaster research: Trends and progress. *Risk, Hazards & Crisis in Public Policy, 12*(4), 374-392.

Wolbers, J. (2022). Understanding distributed sensemaking in crisis management: The case of the Utrecht terrorist attack. *Journal of Contingencies and Crisis Management, 30*(4), 401-411.

World Economic Forum. (2023). *The Global Risks Report 2023*. https://www.weforum.org/publications/global-risks-report-2023/

Wright, A. (2005). The role of scenarios as prospective sensemaking devices. *Management Decision, 43*(1), 86-101.

One System to Connect them All - A Core System for Realizing Integrated Command and Control Research

Björn JE Johansson
Linköping University
bjorn.j.johansson@liu.se

Jesper Tordenlid
Combitech AB
jesper.tordenlid@combitech.se

Jonas Lundberg
Linköping University
jonas.lundberg@liu.se

Jens Alfredson
Saab AB, Aeronautics
jens.alfredson@saabgroup.com

ABSTRACT

Command and Control (C2) for emergency and crisis response is a n important field that poses several research challenges. This paper suggests that Command and Control should be viewed from a holistic point of view, that is the C2 activity system, and that technical C2 support systems must be understood in the broader context of the whole C2 activity system. Studying individual components of C2, such as decision-making, is important, but studying individual components mask out the complexity of C2 as a socio -technical activity. The paper presents an approach to studying C2 based on a generic C2 model combined with a technical core system for connecting different sub-technologies and services – which creates a re-configura ble environment for research efforts in the field of C2 that can span different activity system parts and focus areas.

Keywords

Command and Control, Core system, Command and Control Research, Socio -technical systems

INTRODUCTION

Command and Control (C2) systems are socio-technical systems created with the intent of coordinating resources to manage an unfolding situation. While C2 systems can be fully prepared before a crisis, as seen in military organizations, the degree of preparedness varies in civilian contexts (Trnka & Johansson, 2009). Crisis response without some form of C2 has a small chance of success. C2 systems provide "a foundation for resilience and the management of emergencies and crises" (Landsberg, Schmidt, & Mudimu, 2022, p. 392).

This paper addresses the important issue of enabling holistic evaluations of C2 technologies. While studies of individual C2 support systems, such as decision support or situation awareness technologies, are common, few studies take an overall grasp of the issues of C2, especially when considering novel or prototype technologies that have not yet been integrated into existing C2 support system configurations. While many such evaluations are done with the best of intentions, they are of limited value due to the inherent limitations of the approach. These limitations can include evaluations of individual C2 support systems providing information only about the impact of a particular application and its effect on an individual decision-maker or team, or the study only considering a specific part of C2, such as planning. There are many explanations for this. First, most research or development projects have a limited scope and can only cover a specific part of a C2 system. Second, C2 systems are inherently difficult to evaluate as they often comprise a large number of technologies and individuals who are involved (Landsberg, Gleibs & Mudimu, 2023). Third, the process itself is indirect (this will be discussed further below), and it is hard to differentiate between effectiveness and efficiency (Baroutsi, 2018). However, a very practical issue also haunts C2 research, particularly research involving prototype systems (as is often the case in research):

in many cases, it is hard, if not impossible, to connect all technical systems used in the C2 system. In this work-in-progress paper, we present an effort to create a core system that makes it possible to connect a variety of C2 support systems to enable a holistic approach to C2 in crisis or emergency response, henceforth called the *Core system*. A scenario has been developed within the *WASP Research Arena for Public Safety*[1] (WARA PS) project (Andersson et al., 2021) to capture the challenges of emergency response. This scenario has been used in combination with the Core system to demonstrate the potential of emerging technologies, such as semi-autonomous drones for sea rescue. This case is presented in the paper to show how a core system that connects multiple prototype technologies can be used to study the impact of such technologies on C2. Further, the future development of the Core system is presented with an example of large-scale drone operations. These examples are both part of an overarching project arrangement called Virtual Demo (VD) which focus a number of demonstrator projects around a search and rescue scenario. To establish a firm connect with existing theories on C2, the generic model of C2 developed by Landsberg, Schmidt and Mudimu (2022) is used to show the importance of capturing the diverse aspects of C2.

COMMAND AND CONTROL – WHAT IS IT?

C2 is often considered as a core capability of crisis response (Olsén et al., 2023; Landsberg, Schmidt & Mudimu, 2022; Landsberg, Gleibs & Mudimu, 2023)., 2022), although it is often used interchangeably with "Crisis management" and similar terms. How C2 in crisis management is organized and conducted differs between different nations and even within nations (Sahin, Kapucu & Unlu, 2008). C2 as a research field is inter-disciplinary and has been approached from several scientific fields and theoretical strands (Wikberg, Granåsen & Johansson, 2021; Grant, 2018). Depending on the scientific point of departure, C2 research tend to focus on spe cific aspects of C2, such as decision-making (Brehmer, 2006; Bryant et al., 2004), leadership (Taylor & Rosenbach, 2005; Bullis, 2003; McCann & Pigeau, 1999), situational understanding (Boyd, 1987; Endsley, 1995), systems thinking (Lawson, 1981), adaptive capacity (Alberts & Hayes, 2003), management (Taylor, 1947), organization (Mintzberg, 1979), to mention a few.

Approaches to C2 (NATO, 2010; 2014) range from loosely organized, collaborative, and lacking centralized authority, to strict hierarchical organization with centralized mandate. Independent of how the C2 system is manifested, the purpose of C2 remains the same, to arrange activities and capabilities in time and space to create certain effects in the pursuit of specified goals (Johansson, Herkevall & Lantz, 2023). This implies that C2 comprise the activities of defining goals and coordinating resources to pursue these goals. In practice, this demands several processes to be successful, such as the ability to assess what is happening, dividing tasks and responsibilities, monitor progress, and communicating with executing parts of the concerned organizations. C2 thus consist of a number of activities which form a purposeful system – failure to do so will inevitably result in poor situational understanding, inefficient utilization of available resources, uncoordinated efforts, and possibly increased risk for responding units.

MODELS OF COMMAND AND CONTROL

A *C2 system* is comprised of the involved personnel, the equipment they use, the processes that take place, and the rules or laws that govern work. A C2 system can thus in many respects be seen as equivalent to the definition of a Socio-Technical System (Baxter & Sommerville, 2011 Emery & Trist, 1960). The important point, derived from the use of the word "system", is the holistic perspective of C2 as system created to produce not only orders, but also a system that continuously set goals, assess the situation at hand, create plans, and evaluates the effect of actions. Several attempts to create models of C2 have been made. As in the case of theoretical perspectives on C2, the models differ depending on which discipline they originate from and the purpose of the model. Roughly, a distinction can be made between structural models, that relate different components within the C2 -system, decision models which describe the perception-decision-action cycle, such as Boyd's OODA-loop (Boyd, 1987), and procedural models that describe the various steps of a command cycle in terms of the C2-functions involved and their relation in time, such as Brehmer's Dynamic OODA-loop (Brehmer, 2006; 2009). There are also conceptual models that relate specific dimensions of C2 to each other, most notably the NATO SAS-050 C2 approach space (NATO, 2006). While most of these models were designed in the context of military operations, many of them have been used in the context of crisis management and response due to the lack of established model within the field of crisis management and response. Fortunately, a major research effort has been conducted by Landsberg, Schmidt and Mudimu (2022) in creating a generic model of C2 that can be used in crisis management and response research.

[1] See http://portal.waraps.org/

A Generic Model of Command and Control

Landsberg and his colleagues endeavoured to create a generic C2 model for crisis management with the purpose of supporting research on C2 in terms of evaluation and design (Landsberg, Schmidt & Mudimu, 202 2). The model is based on seven different documents from five different countries describing C2 processes for emergency response and crisis management (ibid). They also used the NATO Code for Best practice for C2 Assessment (NATO, 2002) when designing the generic C2 model to assure that it can be used for evaluation of different aspects of C2. The generic C2 model offers a variety of approaches for research, such as supporting the identification of appropriate key performance indexes (KPIs), identifying important milestones in the C2-process, and identifying relevant components to includes in the analysis (Landsberg, Schmidt & Mudimu, 2022). The model, in its simplified form, includes seven main process steps (denoted A-G, see Figure 1). Landsberg and colleagues aimed to make the model scalable in the sense that it should be applicable to a range of situations from on-site emergency response to large-scale crisis management.

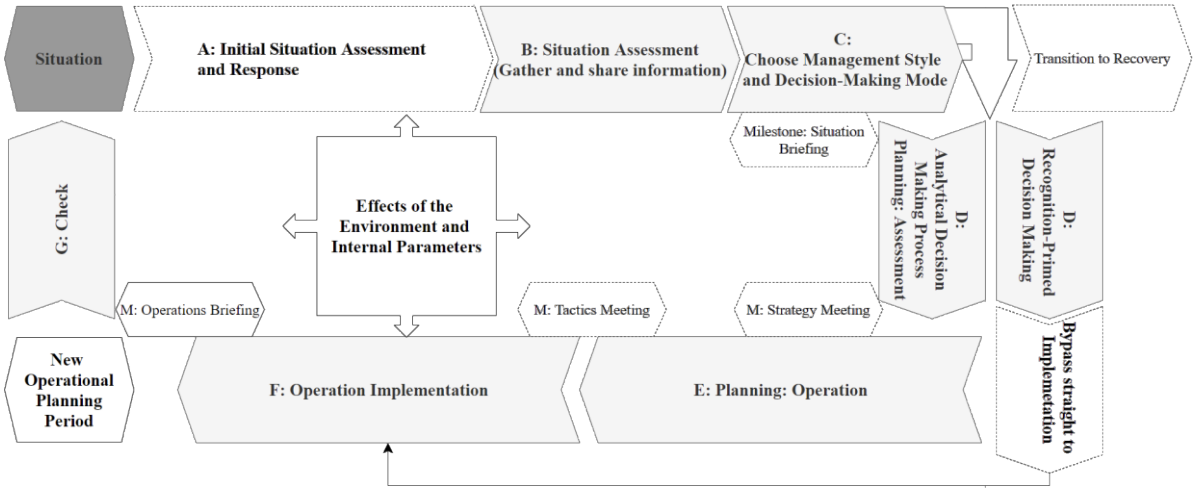


Figure 1. A simplified generic model of C2 (used with permission from the original author, Landsberg, Schmidt & Mudimu, 2022)

In the context of this paper, the model is appealing as it provides a holistic C2 model firmly based on real-world examples and best practice that can be used as a point of departure for identifying and understanding the need for connecting various C2-components and evaluate them as one system rather than isolated parts. The fact that the model is firmly based on a number of real-world C2-systems provides ecological, as well as face, validity. The fact that the model it is relatively scalable in terms of command levels also makes it relevant for many different scenarios, ranging from emergency response to crisis management. However, to make holistic evaluations or research on C2, all or most parts of the model must be in play. This demands at least one of the following things: a full-scale exercise involving all C2 functions that would be in play in a real-world case; a dedicated simulation environment that provides all C2 functions, or, a system capable of connecting a number of individual C2 functions that jointly form a complete C2 system. We argue that the latter approach is the most viable one as this creates a flexible solution that also enables joint research efforts where different research groups can provide solutions that can be developed at different paces and be brought together for more holistic evaluation. Also, such a system should be modular in the sense that it should be possible to exchange individual systems without interfering with the operation of other systems, apart from the fact that certain data generated by that function will not be generated unless it is replaced with another system. Below follows a presentation of such a Core system in the WARA PS system.

ONE SYSTEM TO CONNECT THEM ALL

Similar to the generic C2 model, the Core system of the WARA PS project constitutes a holistic approach, well suited for high level research in C2. As mentioned previously, WARA PS is a Swedish national effort to create a research arena for artificial intelligence (AI), autonomous systems, and future software systems with the aim of shortening knowledge transfer between academia and industry. This collaborative environment comprises test beds that allow for research and development of unmanned surface vehicles, drones, and underwater vehicles (Andersson, et al., 2021). The implementations in the WARA PS Core system maps well to the seven main process steps in the generic model (**Table 1**). The Core system is composed by three interacting layers: the Service layer,

the Tools layer, and the Agents layer (see **Figure 2**) whom all are supported by the Core. The Core part of the system serves as a broker for messaging, packeting, rights, APIs etc. The Tools layer serves as the interface to human users and is essentially where the C2 functions can be realized, and its effects be studied. The Services layer provides necessary services to all the tools, agents and users, as well as other services available. The Agents layer connect both real-world and virtual vehicle, sensors or agents with the other layers and services. Depending on the scenario at hand, different agents, services and human-machine interfaces can be combined. The Core system is therefore flexible enough to support C2 on different levels, ranging from individuals operating equipment to more overarching orchestration of activities on tactical and strategic levels. It is domain agnostic and could be used to connect activities in different domains, such as cyberspace and the physical world. Several activities consisting of individual C2 loops working towards different goals could be set up in a collaborative environment which is transparent and accessible from a research point of view.

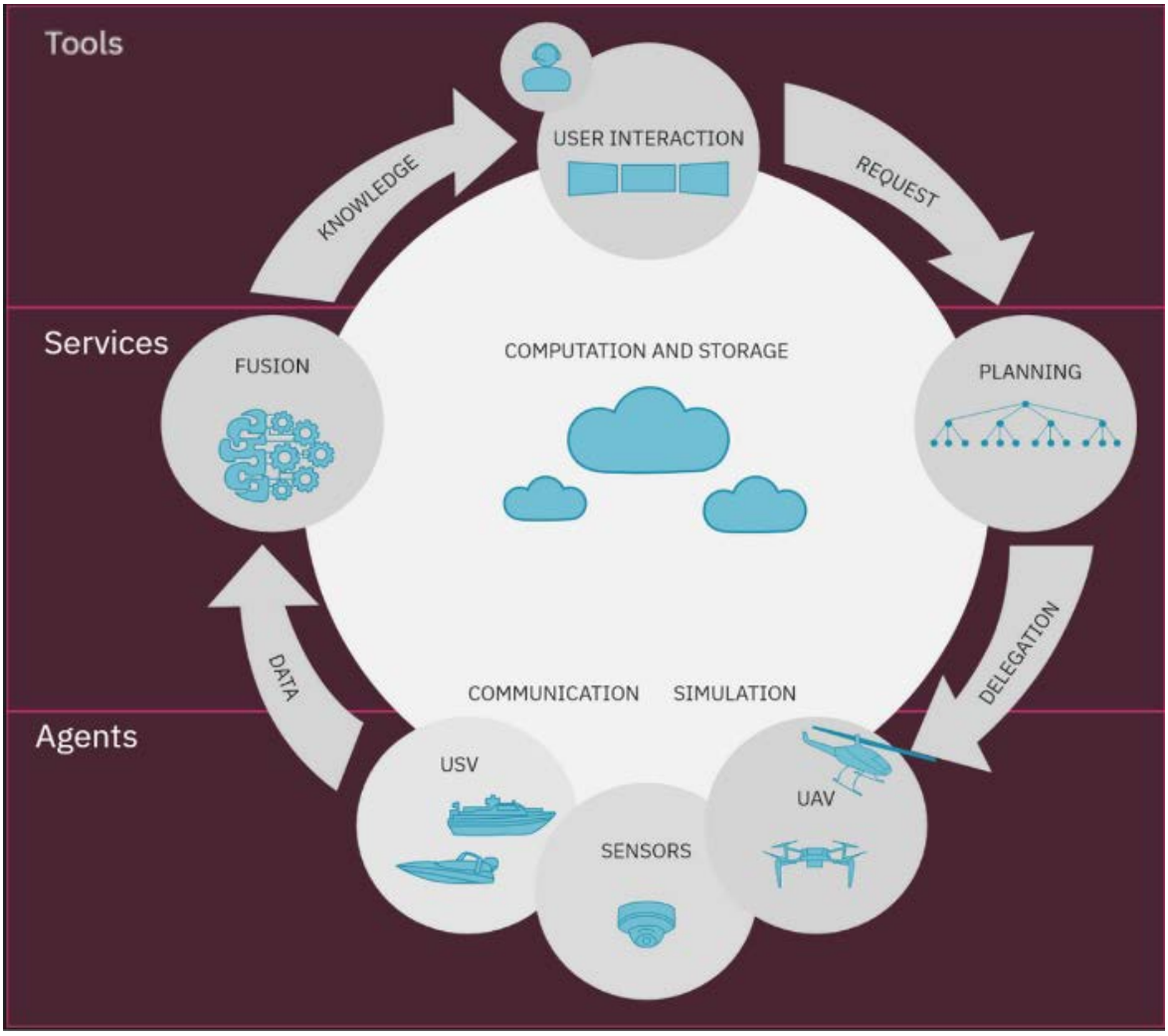


Figure 2. The Core system.

While the Landsberg generic model describes how individual C2 functions are connected through process, the Core system offers a multitude of possible configurations. The Core system can be configured to plan and organize certain aspects of C2, such as planning and analysis of sensor data , both in the form of services provided through the service layer, see **Table 1**. It can also be configured to provide situational pictures to human users, and to study their interaction and behaviour.

Below follows a description of the current status of the WARA PS Core system based on a demonstration of the system made in 2023 the *WARA PS Search and Rescue demonstrations*. This is followed by a description of a system that will be connected to the WARA PS Core system in the future, the *UTM City Drone simulation*, which will add a strategic level of operations to the overarching C2 system once incorporated with the Core system.

Table 1. Mapping the generic model to the Core system of the WARA PS case.

Generic C2 model	Core system – Examples
A. Initial Assessment and Response	Agent layer – Agents interacting with the environment or simulated environment to collect and analyze data Service layer – anomaly and object detection on fused data to assess situation Tools layer – Brief initial Situational Picture Looped until initial situation assessed
B. Situation Assessment	Agent layer – sensor data from Agents Service layer – sensor fusion and (re)planning Tools layer – Situational Picture(s) Looped until situation assessed
C. Chose Management Style	Human Activity – Commanders must decide if they are to manage the situation based on trust and dissemination of decision rights to sub-ordinate units, or if the situation demands a more centralized, directive command style. Combinations of the approaches may occur depending on level of command and domain. Agent layer – what kind of C2 management does the agents demand? Can they be configured to act upon different styles. Service layers – what kind of structural configurations of the C2 system is possible. The Core system offers a high degree of flexibility. Tools layer – what management styles are supported by the HMI components? For example: must orders be produced in the form of directives, or does the system offer some flexibility in the formulation of orders?
D. Decision-Making	Tools layer – Decision support. Decide complexity of operation and needed amount of planning
E. Planning Operations	Tool layer – HMI for human interaction Service layer – planning of operations based on available resources and operational needs Agent layer – provide information of available agents, capabilities and services.
F. Operation Implementation	Tool layer – C2 command systems Service layer – Fusion services for analysis like object detection, anomaly detection and data fusion. Also (re)planning Agent layer – execution of orders, simulation, search, move, pickup and drop objects. Provide data for analysis.
G. Check	Tools layer: Decide if new strategy needed or operational objectives fulfilled. Services layer: Support for analysis and evaluation of operation progress.

Current Status - The WARA PS Search and Rescue Demonstrations 2023

The WARA-PS Core system has been used during demonstrations of search and rescue missions, using real and simulated unmanned arial vehicles (UAVs), unmanned ground vehicles (UGVs) and unmanned surface vechicles (USVs) together with human agents and humans in C2 . The unmanned vehicles had varying degrees of autonomous capabilities and act as agents integrated in the system. They communicate using high level commands such as Search Area, Move To, Pick up and Drop. Each vehicle acts independently and interprets the commands/tasks in its context and executes based on its capabilities (see **Figure *3***). The system offers a way to communicate and share tasks and data in the same mission.

A scenario has been used in the demonstrations. In the scenario, victims of a fictive boat accident were spread out in water and on land. Victims in water were represented by floating tennis balls to challenge detection algorithms and create a fast and interesting demonstration illustrating activities and decisions. On land the victims were played by humans acting in the role as injured.

A simplified activity model of the operation conducted in the scenario was used, where the autonomous agents conducted the following activities simultaneously to search and rescue victims in the scenario. The involved agents changed activity depending on the type of task in collaboration with human decisions.

- **Prepare** for a probable upcoming situation, by placing drones and aid on right place
- **Surveillance** to detect anomalies, as an ongoing activity to identify changes or assess situation
- **Search** to detect interesting objects as an activity to assess (initial) situation
- **Inspect** detected objects to identify them as victims as candidates to rescue
- **Rescue** by delivering useful aid to identified and prioritized victims
- **Evaluate** to learn, improve and be ready for a new planning period or a new operation

In an operation all those activities are ongoing as parallel activities and agents shifts activity based on needs to perform an efficient operation.

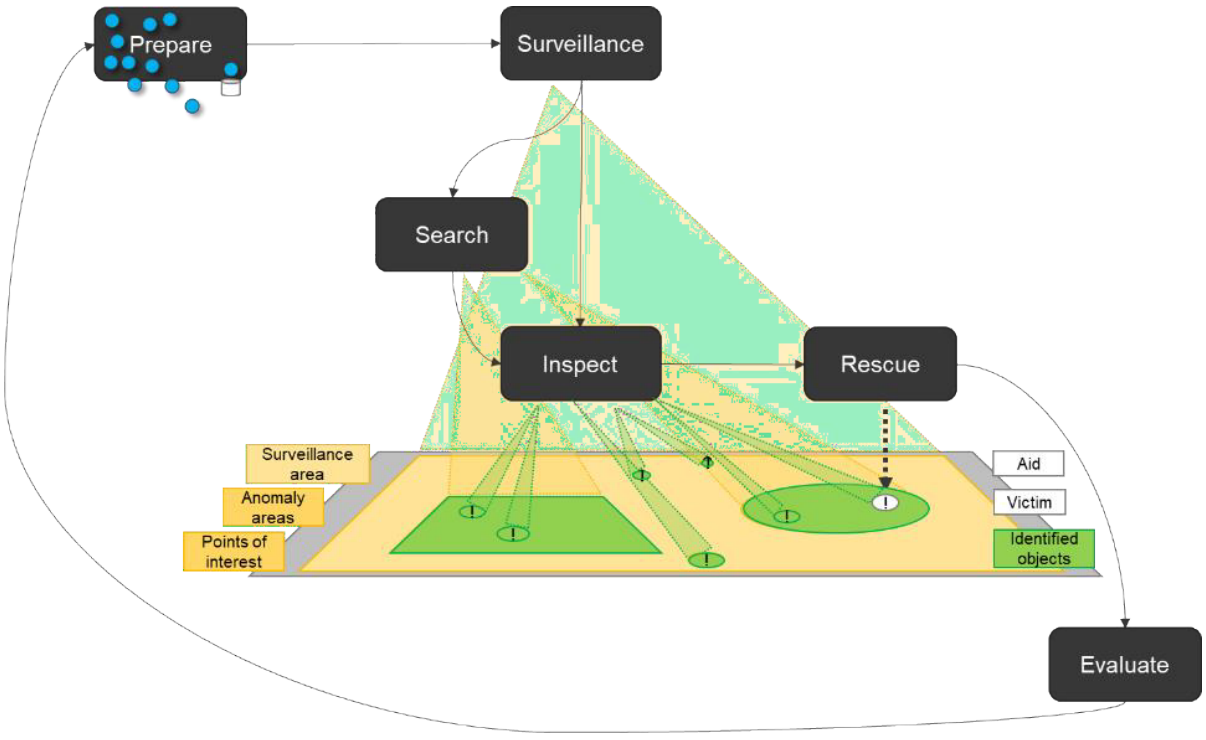


Figure 3. Activities in a search and rescue operation performed by using the WARA-PS core system with autonomous agents.

The demonstrations were performed in September 2023 in Sweden at the Gränsö drone testing range during the WARA-PS demonstration week. With an audience of over 200 persons the scenario was played multiple times on two days where the activities constantly changed due to weather and different introduced disturbances.

During the demonstrations the physical agents was represented by:

- Two UAVs flying over the scene with cameras assessing the situation performing surveillance and search for victims. To rescue they have a device to drop medical kits at safe distance from victims. The UAVs was prepared to support both activities over land and water based on upcoming needs by changes in the situation.
- One Boston Dynamics Spot robot for close inspection of detected objects. The Spot robot can also pick up the dropped medical kits and move close to victims for rescue.
- One Saab Piraya USV boat to search the sea using camera and lidar. This larger USV was meanwhile performing autonomous navigation with collision avoidance testing.
- Four small mini-USVs working together to inspect detected objects in water by moving closer to the objects detected by the Piraya USV and UAVs. Using cameras, the mini-USV supported the humans in C2 to identify and prioritize the objects as victims to be rescued.
- A human on stand-up paddle board equipped with a mobile phone application to serve with capability of moving close to victims and pick them up on the board for transport to shore.

The core system offered exchange of data and tasks to provide all agents and humans in C2 with shared situational picture including objects, victims and changes. The C2 tools used was the WARA PS *Arena Map* where two human operators supervised the operation and supported with decisions by commanding the agents. Additional instances of the C2 were active to give the audience a situational picture containing live video feeds, map and 3D overview. The USVs had their own USV C2 to supervise and make decisions on lower level to support the larger USVs (Piraya).

Communication in the system was a mix between 4G/5G, wifi and dedicated radio links. Human operators in C2, field agent on paddle board, safety pilots for the vehicles and demonstration leader used separate voice communication. Services developed from different organizations was executed in cloud and on agents to support object detection, anomaly detection and task delegation/distribution. Besides the physical agents the system introduced multiple simulated agents to investigate how the system and operation can be scaled.

There was also a couple of other separate demonstrations ongoing simultaneously using air space and facilities that was coordinated to give audience an insight of how the core system and resources can be used and what has been done during past year. Collected data as positions, decisions and video was recorded for analysis and evaluation. Besides the sha red recording many of the involved systems has their own storage of information to be extracted and used for research purposes and to improve she systems.

Ongoing Development – Integrating the UTM City Drone Simulation at the Strategic C2 Level

An example of adding new integrated functionality, to combine lower levels of C2 with higher levels, in the WARA PS systems is the *Unmanned traffic management, for air traffic in cities* (UTM City) platform (Westberg, Palmerius & Lundberg, 2022; Lundberg, Arvola & Palmerius, 2021). The Core system contains a mission-management component, the *Arena Map*, where resource usage can be planned, as well as specific routes to use for drones in the air, at sea, on land. UTM City adds capabilities for coord inating the use of physical air space between co-located missions. The UTM CITY platform is a collection of software for drone traffic simulation and control. It is the result of eight years of research on drone traffic planning and management, visualization, and human-automation interfaces. The platform consists of two main parts: a drone traffic infrastructure simulator that we call *dronesim*, and a user interface for this simulator.

The *dronesim* module handles the simulation of all activities involved in drone traffic, and has been designed to test separation of concerns and communication between the actors by isolating the simulation of 1) drone operators, setting up drone activities, sending plans for authorization and controlling the drone flight, 2) airspace authority, holding the airspace design, responding to queries about traffic situation, and validating, modifying and authorizing flight plans, and 3) activities performed by a unmanned tra ffic control officer (UTCO).

The *user interface* module connected to this simulation module provides a 3D map over the designated area and simple means to activate various data and layers, and two different menus for scenario control and air traffic management, respectively. The menu for scenario control provides means to draw activities for dron e operators on the map, for automatically generating potentially hundreds of drones quickly and with little effort. The menu for air traffic management lets the user control airspace parameters, such as minimum allowed distances, as well as create, activate/deactivate and adjust no-fly-zones.

The dronesim module automatically reacts to airspace changes, sending signals to the affected operators that their plans need to be renegotiated. They will subsequently calculate new plans based on the remainder of their ongoing activities and submit these for authorization. The airspace authority uses the airspace configuration to adjust these

plans if necessary and responds with the updated plans. The UTM CITY platform also has a streaming log module, to capture what goes on in the system, for real-time analysis or after-action review.

By the integration of this functionality in the WARA PS case, yet another concrete implementation related to the generic model is available for the study of C2, especially concerning drone traffic planning and management. Not only are such studies supported by the implemented functionality, but also by the C2 contextual framework of the Core system of the WARA PS case, see **Figure *2***, and **Table *1***.

In a simulation, UTM CITY operators can work on a more strategic level than when using the Mission Management module in the Core System. In UTM CITY, common airspace structures are established, such as flight common-use routes or airspace areas for shared operational spaces, as well as flight levels, geofences and geocages for isolation of operation. In the Mission Management component, the Arena Map of the Core System, different agents can request routes or destinations or geofences, that are then coordinated in UTM CITY. UTM CITY can also perform basic simulations of prototype services such as mass-deliveries and surveillance, point-to-point deliveries – to be able to identify and foresee problems in the airspace by stressing and simulating situations rapidly (but not on a detailed level). By connecting the sub-systems in the Core System, mission management can be studied at the same time as airspa ce management, in an integrated Virtual Demonstration; or a Physical-virtual Demonstration, or a purely physical demonstration.

Thus, in a setting such as a search-and-rescue mission, different teams need to collaborate and act together. The use of airspace for their missions can then be coordinated in the UTM CITY platform. In future versions of UTM CITY, surface mission areas or sub-surface mission areas could also be managed in similar ways.

DISCUSSION – CREATING A HOLISTIC C2 RESEARCH ENVIRONMENT

Creating conditions for holistic research on socio-technical C2 activity systems is possible. Far too often, aspects of C2 are studied, such as individual ICT components supporting situational pictures, communications, social media analysis etc. While these efforts are both necessary and admirable, they are also problematic from a systems point of view. When studying "bits and pieces", we mask out some of the actual complexity of a C2 system , both in terms of the complexity of interconnected technical systems and in terms of inter-personal aspects of C2. It should also be noted that creating a system like the Core system is not done to impose control over several interconnected systems – it is done to be able to study how C2 manifests and evolves in different scenarios under different configurations of C2 capabilities. The focus is thus on interconnectedness rather than control, on the ability to farm data from a multitude of systems, and on the possibility to study C2 as a whole rather than individual components.

We have exemplified how the integration of a new component in the Core System can enable a new kind of integrated study of C2. Rather than starting from scratch to study collaborative mission-management and operational space management, the Core System integrates two separate m odules (Mission Management and Operational Space Management). By doing so, it is also integrated into the larger whole of the Core System, to make use of other components, and enabling other kinds of studies. For instance, to include studies of autonomous systems operation across mission management and operational space management. While this example focuses on drone operations in emergency response, this must not be the case. The Core system has general capabilities to connect other kinds of systems. As can be seen in **Table *1***, the word "agent" must not be equivalent to a drone, and services and tools must not be specific to drone operations. In this sense, using the generic model has expanded the focus of the Core system by looking beyond demonstrations of emergency response with drones to a more holistic view of experimentation and research on C2 systems. For example, the mission management system, which can be used on the operative level for A. Initial Assessment and Response, B. Situational Assessment, and F. Operational Implementation, can be combined with the UTM City system to include D. Decision Making and E. Planning Operations on the strategic level. Further, these systems were developed independent of each other, now coming together under the Core system umbrella .

Bringing Research Into the Picture

The WARA PS Core system example demonstrates how a usable research tool for C2 can represent a concrete representation of relevance to the generic C2 model, and at the same time allow for variations in implementation of relevance for applied C2 research. Therefore, the Core system is also well suited for studies of the ma in processes, their interdependence, and their internal characteristics. Indeed, the Core system has been described as "a system architecture for industrial and academic collaboration" (Andersson et al., 2021). On the one hand, adding of components can enable operational capabilities, to carry out exercises. On the other hand, can also add research capabilities such as logging and recording of exercises. When integrating components, it is important to

integrate both kinds of capabilities, to enable studies that go across different C2 activities. As pointed out by Landsberg, Gleibs and Mudimu (2023) in their systemic approach to C2 assessment, C2 can, and possibly must, be evaluated on several levels. They suggest that C2 systems can be evaluated in terms of *physical characteristics*, *structures and processes*, and *C2 system effectiveness*. The Core system focus mainly on creating conditions for evaluating C2 system effectiveness but can naturally be used to examine the other levels as well. The modularity of the system may create some challenges in this regard as a joint analysis of a multitude of supporting systems and services on all core system levels (Tools, Services, Agents) must be conducted to understand the physical level, which the systemic approach of Landsberg and colleagues describe in terms of bandwidth, connectivity, and computer capacity. However, structure and process, described in terms of for example time to complete tasks, time to transfer information, number of personnel involved in different activities etc (Landsberg, Gleibs & Mudimu, 2023) could easily be captures by comparing different parts of an implemented scenario with different inputs to, and interactions with, Core system components. The third level, C2 system effectiveness is in the Landsberg and colleagues framework proposed to evaluate situational awareness, workload, sensemaking, synchronization of effort, communication, and leadership (ibid.). These aspects of C2 are also possible to evaluate in the context of the Core system, but we suggest that focusing on scenario-based measures of success could be an addition. Measuring human factors aspects of C2 such as situational awareness and workload is interesting if we want to understand how different C2 components impact humans in the system, but we also need to keep focus on the overarching performance of the C2 system and evaluate the effects on human operators in the light of this.

The WARA PS core system is dynamically adjustable for applied studies of C2. This means that a wide range of research questions can be, and has been studied, of which some examples has been provided above. The way the WARA PS scenario and the implementation in the core system has been designed, makes it not only suitable for system research, but also for research on methods and approaches. Since also future capabilities, not yet realised in today's practise is represented, it is well suited for applied research of future challenges and possibilities in C2. The main goal for the future development of the Core system is therefore to create possibilities for connecting further, and future, systems to the environment.

CONCLUSIONS

To create possibilities for holistic evaluations of socio-technical C2 activity systems, it is important to not only be able to connect different C2 solutions, but also that capabilities for studying C2 are in place. The Core system is based on modularity and interconnectedness rather than a tailor-made approach for studying a specific problem. Instead, it can be configured to enable applied research on a number of issues related to C2 in emergency and crisis response. A case describing such research, as well as suggestions for future development has been described above in terms of the WARA PS Gränsö demonstrations and the UTM City simulator. The generic C2 model by Landsberg, Schmidt and Mudimu (2022) has been used to show how the layers of the Core system can be connected to a model of C2 to assure that all relevant aspects of C2 are captured. Using such a model of C2 is necessary to assure that no aspects of C2 are neglected. Further, the generic model can be used to understand what parts of C2 a newly developed technology can support, and how it is related to other C2 systems, i.e. what types of connectivity it must have to be part of the whole. The generic C2 model also provide key performance indicators that are useful for evaluating the impact of different kinds of C2 support systems or agents, for example in the form of drones or new autonomous capabilities.

REFERENCES

Alberts, D. S. & Hayes, R. E. (2003). *Power to the edge: Command, control in the information age*. Washington, DC: CCRP publication series.

Andersson, O., Doherty, P., Lager, M., Lindh, J-O., Persson, L., Topp, E.A., Tordenlind, J., Wahlberg, B. (2021). WARA-PS: A research arena for public safety demonstrations and autonomous collaborative rescue robotics experimentation. *Autonomous Intelligent Systems.* 1(9), 1-31.

Baroutsi, N. (2018). A practitioners guide for C2 evaluations: Quantitative measurements of performance and effectiveness. *Proceedings of ISCRAM 2018*, Rochester, May 20-23.

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with computers*, *23*(1), 4-17.

Boyd, J. (1987). *A discourse on winning and losing.* Maxwell Air Force Base, AL: Air University Library Document No. M-U 43947 (Briefing slides).

Brehmer, B. (2006). One loop to rule them all. *Proceedings of the 11nd Int. Command and Control Research and Technology Symposium (ICCRTS)*, Cambridge, MA, June 20-22.

Brehmer, B. (2009). From function to form in the design of C2 systems. *Proceedings of the Int. Command and Control Research and Technology Symposium (ICCRTS)*, Washington, D.C. June 15–17.

Bryant, D. J. (2006). Rethinking OODA: Toward a modern cognitive framework of command decision making. *Military Psychology*, 18(3), 183-206.

Bullis, C. (2003, May–June). Developing the professional Army officer: Implications for organizational leaders. *Military Review*, 57–62.

Emery, F. E., & Trist, E. L. (1960). Socio-technical systems. *Management science, models and techniques*, *2*, 83-97.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37, 32–64.

Grant, T. (2018). Common topics in C2 doctrine for emergency management. *Proceedings of ISCRAM Asia Pacific*, Wellington, New Zealand, November 5-7.

Johansson, B., Herkevall, J. & Lantz, M. (2023). A Framework for Understanding the Relationship Between C2 Theories, Military Long-Term Planning, and Empirical Studies. *Proceedings of the 28th International Command and Control Research and Technology Symposium (ICCRTS)*, Laurel, November 28-30.

Landsberg, L., Schmidt, J., & Mudimu, O. A. (2022). Synthesising Comparisons to Develop a Generic Command and Control System. *Procedings* of *ISCRAM 2022*, Tarbes, May 22-25.

Landsberg, L., Gleibs, T., & Mudimu, O. A. Design of a Systems Theory Approach for the Evaluation of C2 - Systems. *Proceedings of ISCRAM 2023*, Omaha, May 29-31.

Lawson, J. (1981). Command control as a process. *IEEE Control Systems Magazine*, *1*(1), 5-11.

Lundberg, J., Arvola, M., Palmerius, K.L., (2021). Human Autonomy in Future Drone Traffic: Joint Human–AI Control in Temporal Cognitive Work. *Frontiers in Artificial Intelligence* 4:704082.

McCann, C., & Pigeau, R. (1999). Clarifying the Concepts of Control and of Command. *Proceedings of the 1999 Command and Control Research and Technology Symposium (ICCRTS)*, Rhode Island, June 29 – July 1.

Mintzberg, H. (1979). *The Structuring of Organizations.* Prentice Hall. Englewood Cliffs, NJ.

NATO (2002). *Code of Best Practice for C2 Assessment*. Dod Command and Control Research Program.

NATO STO SAS-050 (2006). *Exploring new command and control concepts and capabilities.* NATO Science and Technology Organisation.

NATO STO SAS-065. (2010). *NATO NEC C2 Maturity Model*. NATO Science and Technology Organisation.

NATO STO SAS-085 (2014). *Final Report on C2 Agility*. Brussels: NATO Science and Technology Organisation.

Olsen, M., Oskarsson, P. A., Hallberg, N., Granåsen, M., & Nordström, J. (2023). Exploring collaborative crisis management: a model of essential capabilities. *Safety science*, *162*, 106092.

Sahin, B., Kapucu, N., & Unlu, A. (2008). Perspectives on Crisis Management in European Union Countries:

United Kingdom, Spain and Germany. *European Journal of Economic and Political Studies*, 1(1), 19-45.

Taylor, F. W. (1947). *Scientific Management*, Harper and Row. New York.

Taylor, R. L., & Rosenbach, W. E. (2005). *Military leadership: In pursuit of excellence*. Cambridge, MA: Westview Press

Trnka, J., & Johansson, B. (2009). Collaborative command and control practice: adaptation, self -regulation and supporting behavior. *International Journal of Information Systems for Crisis Response and Management (IJISCRAM), 1*(2), 47-67.

Westberg, J.J., Palmerius, K.L., Lundberg, J., (2022). UTM City—Visualization of Unmanned Aerial Vehicles. *IEEE Computer Graphics and Applications*, 42, 84-89

Wikberg, P., Granåsen, M. & Johansson, B. (2021). Perspectives on Command and Control: Implications for Capability Development and Assessment. In *Proceedings of the 26th International Command and Control Research and Technology Symposium (ICCRTS),* Quebec, October 18-22.

Towards training network resilience to maintain disaster recovery expertise

Paul Savignac
Centre de Génie Industriel, IMT MINES ALBI, Universite´ de Toulouse
paul.savignac@mines-albi.fr

Christopher Zobel
Dept. of Business Information Technology, Virginia Tech
czobel@vt.edu

Eva Petitdemange
Centre de Génie Industriel, IMT MINES ALBI, Universite´ de Toulouse
eva.petitdemange@mines-albi.fr

Nicolas Cellier
Centre de Génie Industriel, IMT MINES ALBI, Universite´ de Toulouse
nicolas.cellier@mines-albi.fr

Robert Weiss
Department of Geosciences, Virginia Tech
weiszr@vt.edu

Matthieu Lauras
Centre de Génie Industriel, IMT MINES ALBI, Universite´ de Toulouse
matthieu.lauras@mines-albi.fr

ABSTRACT

Effective disaster recovery depends on highly-skilled practitioners. Like any sector, however, the education and training systems that build these skills can experience crises that disrupt their ability to function effectively. These can be one-off events (e.g. a major flood) or extend over long periods (e.g. a pandemic or a vocational crisis). Based on the observation that there is a shortage of trainers and teachers in the education sector, as shown by indicator D7 (OECD, 2023), in this article we study the ongoing ability of a French vocational training organization to transmit skills that are useful to society in the event of a natural disaster, and in particular those that will shorten the recovery time of buildings and infrastructures. We propose a model that simulates the flow of people through the educational system and the institution's ability to generate enough trainers to maintain a chain of know-how transmission. In the simulation of this work in progress, we have parameterized the actual success rates, which we vary in order to define a minimum success rate limit to ensure the renewal of human capital. Based on the proposed model, further testing and analysis will be used to characterize a range of resilient flow behaviors and to provide recommendations for maintaining the condition of the human know-how transmission chain.

Keywords

Resilience, Education System, Evaluation, Key Performance Indicator, Decision Support, Steering

INTRODUCTION

The impacts of climate change on modern societies are still in their early stages. An acceleration in the frequency and magnitude of natural disasters will increasingly lead to damage to people and property (Intergovernmental Panel On Climate Change, 2023). Beyond the public policies implemented to address these crises, it is crucial to recognize the need for skilled workers to help restore, as quickly as possible, the infrastructure, buildings, and industrial facilities essential for the proper functioning of human societies (Bruneau & Reinhorn, 2006). The capabilities required of these workers, in terms of know-how, emotional intelligence, and an understanding of the most effective actions available, take many years to develop and require significant ongoing investment in order to avoid obsolescence (Sandborn & Prabhakar, 2015).

Many of the decision-making skills needed by such expert tradespeople are also needed for effective decision making in other crisis management contexts. For example, it is important to have the capacity to understand complex

situations, to withstand stress and pressure, and to make quick decisions, as well as to coordinate and communicate, to demonstrate adaptability, ethics, and responsibility, and to synthesize feedback. These skills therefore represent a critical resource that needs to be developed and maintained in order to avoid a shortage of expertise that could have devastating repercussions in a crisis.

In France, the *Association Ouvrie`re des Compagnons du Devoir et du Tour de France* (AOCDTF) has been helping to train high-level experts in over thirty trades, mainly in construction and industrial technology, for several centuries. This type of manual know-how enables emergency repairs to be carried out and essential infrastructure to be maintained, such as public buildings, power plants and housing. In so doing, it contributes to the resilience of society.

Here we consider resilience as *"the ability to prepare and plan for, absorb, recover from or adapt more successfully to actual or potential adverse events"* (National Research Council, 2012). The expertise provided by crisis response and recovery organizations is critical for helping society to be resilient, but these organizations may also face disruptions to their own ability to properly train new workers. To play their part in community resilience (Valinejad et al., 2022), therefore, organizations like the AOCDTF (or the *"Compagnons"*) must also have resilient education and training systems, in order to ensure that each organization maintains its own human chain of know-how. This resilience means taking into account the flow of people through the system, and in particular ensuring that the system generates sufficient numbers of graduates and has the capacity to continue doing so. In the case of the *"Compagnons"*organization, this means ensuring that enough young professionals become trainers, so as to maintain the training capacity for the incoming apprentices. When this isn't the case, it creates problems for the organization's economic model, which is then forced to call on former members of the system or external professionals whose remuneration is higher. Furthermore, the pedagogical model of the *"Compagnons"* is based on community life and on there being only a small age gap between learner and supervisor. A disruption in the training system thus also puts pressure on young people who have not yet completed their apprenticeship to take on the job prematurely to fill the gap, which is an undesirable outcome.

With this in mind, this article seeks to contribute to improving the resilience of human capital in training organizations, focusing particularly on the issue of the shortage of trainers for the highly skilled workers needed for disaster recovery efforts. We propose to model the flow of learners and its consequences on the flow of teachers or trainers, in order to better understand the resilience of these educational organizations to internal disruptions, and thus to support decision-makers in their efforts to respond to and recover from such disruptions more effectively. We first introduce a generalized model and provide its mathematical formulation. We then discuss the specific context of the *"Compagnons"* organization in more detail, and apply the model in that specific context, to illustrate its use.

BACKGROUND

Educational systems are intricate organizations, challenging to comprehend to the extent that advanced tools are necessary for modeling them to anticipate their behaviors (Jacobson et al., 2019). Furthermore, such systems are challenging to transform (Burner, 2018), and they often operate in a volatile, uncertain, complex, and ambiguous (VUCA) environment (Waller et al., 2019). This environment exposes such systems to various risks that they must be capable of assessing in order to devise strategies to ensure their proper functioning (Dehdashti et al., 2020). In the face of such complexity, asking questions about resilience such as Meerow et al. (2016)'s five W's (Who? What? When? Where? Why?) is important for allowing educational systems and organizations to gain a broader understanding of the specific context within which they seek to achieve such resilience.

With this in mind, Table 1 enumerates the risks impacting educational systems, classifying them according to the transdisciplinary perspective required for the integration of sustainability and resilience into risk (Nielsen & Faber, 2021). It is self-evident that these risks are not exhaustive, and each educational organization must undertake a comprehensive survey tailored to its environment.

Our focus here is on the trend towards a shortage of teachers within such systems, which is a challenging phenomenon because it unfolds over an extended period of time. It offers an example of a so-called "slow-motion disaster" (or "slow-onset disaster") where the impacts (in this case, a corresponding shortage of highly skilled workers) are not immediately felt. Pierrehumbert (2006), in discussing climate change and carbon dioxide emissions, defines slow-motion disasters as events where the consequences manifest over a century or more. Draus (2009), in the study of Detroit's population and drawing on numerous other works, refers to slow-motion disasters by focusing on the consequences of multiple vulnerabilities that can accumulate and whose effects become visible over the long term. He introduces a sociological component, represented by a cascade of cause-and-effect relationships. Additionally, he identifies slow-motion disasters at the structural, process, or situational state levels.

Table 1. Risks in Educational Organizations

Risk Family	Risk	Capacity from / to
Human Assets	Lack of trainers	Managing human resources
	Excessive attrition rate	Supporting learners
	Demographics	Adapting the organization
Production	Teaching obsolescence	Update skills
	Reputation	communicate
	Cheating	Fostering integrity
	Operational under-performance	Leading and empowering staff
Health	Psycho-social	Taking care of people
	Contagiousness	Organizing health constraints
Earth system, Environment, Food, Water, Natural resources	Ethics	Training stakeholders
	Environmental	Minimizing the organization's footprint
	Shortages	Securing supply
Economy	Budget limits	Manage fluctuations
	Corruption	Control integrity
Systems engineering, Infrastructures, Free market systems, Governance	Universal access uncertainty	Plan
	Organizational under-performance	Continuous improvement
	Technological disruption	Adapt strategy
	Administration	Organize and plan for the future
	Competition and globalization	Satisfy stakeholders
	Legal	Track activities
	Regulatory	Ensure compliance
	Crisis management	Prepare for action

The evaluation of educational systems is most often used for assessment rather than for guidance (Michel, 2000), and there can be challenges with accepting evaluation by those involved in vocational education (Jorro, 2007). Our ambition is to provide indicators that can contribute to both measuring and improving the resilience of human trainer resources in this context. We thus need to know inputs, outputs, objectives, and instructions, as well as human, technical, and financial resources, in order to appropriately characterize how such an organization functions. To evaluate the behavior of these processes, we then also need to have key performance indicators (KPIs) at our disposal that will enable us to measure effectiveness (which is the relationship between objectives and result), efficiency (which is the relationship between means and results), and relevance (which is the relationship between means and objectives) (Nadoveza Jelic et al., 2018; Saikia & Bezborah, 2014).

MODEL

To support evaluating the resilience of educational systems in the context of slow-motion disasters, we propose a discrete, generalizable mathematical model that represents the flow of individuals—learners, graduates, and trainers—through time.

By creating this model, we aim to reach two goals: to have a better overall understanding of the dynamics of the system and to be able to answer questions such as the following, in order to guide resilience-focused decision making:

- What is the sustainable growth level of the system?
- What happens if the rate at which learners become trainers varies upward or downward?
- What occurs if the inflow rate slows down (due to demographic changes, different choices of orientation by entrants, etc.)?
- What happens in the event of a crisis (i.e., a sudden decrease in learners or trainers)?

Because of the length restrictions of a work-in-progress paper such as this, our discussion below focuses on only one such question, illustrating the potential of the new mathematical model by analyzing the effect of changes in the student attrition rate.

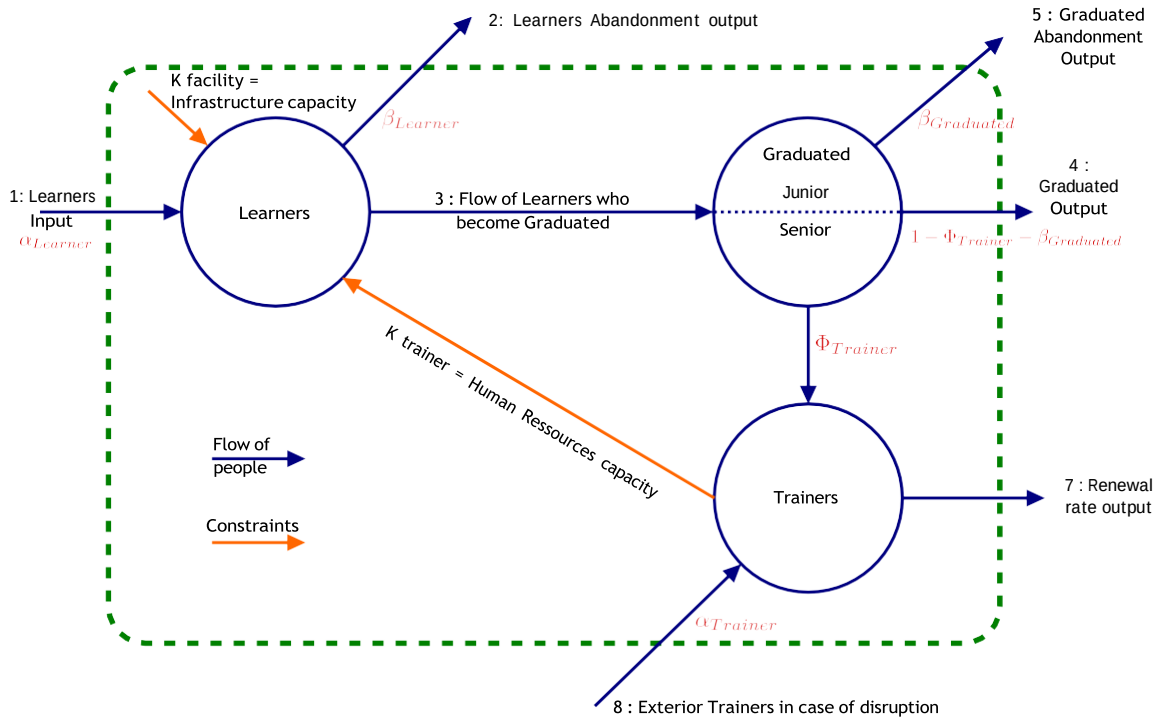


Figure 1. Trainer flow model

In order to enable an education system's managers to project their available human resources (i.e., the personnel who serve as trainers), we propose the model shown in Figure 1. In this model there are 3 pools of resources :

- The Learners : The behavior of the Learners in the system is defined by equation (1). The training will last a duration of $\tau_{Learner}$ years. Each year the pool of learners will suffer losses $\beta^t_{Lear\ ner}$ due to drop-out, reorientation, or breach of contract between the learner and the host company (flow 2). Those who will succeed through all the training will be the input of the Graduated pool.
- The Graduated : The behavior of the Graduated in the system is defined by equation (2) and equation (3). A Graduated is first a Junior for one year before becoming a Senior. These graduates may subsequently use their expertise to support the training efforts for an additional period of time, with this group also suffering $\beta_{Graduated}$ losses due to abandonment and other reasons (flow 5). After a certain time, τ_{Senior} , a compliant flow out of the system takes place, for reasons of career development, updating of knowledge, or other reasons (flow 4). A proportion $\Phi^t_{Tr\ ainer}$ of Senior graduates will then change state to become the input flow of the Trainers pool.
- The Trainers : The behavior of the Trainers in the system is defined by equation (4). Trainers will stay for time $\tau_{Trainer}$ before leaving the system (flow 7). In the event of a shortage of trainers, it is possible to call on external sources of trainers: $\alpha^t_{Tr\ ainer}$ (flow 8).

We also have added several infrastructure constraints: $K_{Facilit\ y}$ is the capacity of the training centers, representing the upper limit of learner intake, and $K_{Trainers}$ is the ideal number of trainers needed for the system to perform as expected. Based on this description of the training system, we introduce the associated mathematical flow model. The relevant parameters on which the flow model is based are as follows :

- τ_{Learner} : the duration of the training.
- τ_{Senior} : the minimum time before leaving the system as a Graduate.
- $\tau_{Graduated}$: the minimum time before becoming a Trainer.
- $\tau_{Trainer}$: the minimum time before leaving the training system.
- $\Phi^t_{Tr\ ainer}$: the Trainer transformation rate: the portion of the Graduated that will become Trainers
- $\beta^t_{Lear\ ner}$: the Learner abandonment rate
- $\beta_{Graduated}$: the Graduated abandonment rate, before becoming either a Senior or a Trainer

- $\alpha_{Learners}$: the input of Learners each year.
- $\alpha^t_{Trainer}$: the input of external Trainers each year.

$$Learner^t = \alpha^t_{Learner}$$
$$Learner^t_i = Learner^t_{i-1}\left(1 - \beta^t_{Learner}\right) \quad \forall i \in [2, \tau_{Learner}] \tag{1}$$

$$Graduated^t = Learner^t_1 \left(1 - \beta^t_{Learner}\right)^{\tau_{Learner}}$$
$$Graduated^t_i = Graduated^t_{i-1}\left(1 - \beta_{Graduated}\right) \quad \forall i \in [2, \tau_{Graduated}] \tag{2}$$

$$Senior^t = Graduated_{\tau_{Graduated}}\left(1 - \beta_{Graduated} - \Phi^t_{Trainer}\right)$$
$$Senior^t_i = Senior^t_{i-1}\left(1 - \beta_{Graduated}\right) \quad \forall i \in [2, \tau_{Senior}] \tag{3}$$

$$Trainer^t = Graduated_{\tau_{Graduated}}\Phi^t_{Trainer} + \alpha_{Trainer}$$
$$Trainer^t_i = Trainer^t_{i-1} \quad \forall i \in [2, \tau_{Trainer}] \tag{4}$$

The parameters $\Phi^t_{Trainer}$, $\beta^t_{Learner}$ and $\alpha^t_{Trainer}$ are time dependent and will vary depending on the system regime. Three types of regime are defined in equation (5) and differentiated by the learner supervision rate as follows:

- Regime 1, which incorporates over-capacity utilization of the organization (one trainer for more than X2 learners). This situation will lead to a deterioration in service in terms of quality of welcome, supervision, and ultimately success.
- Regime 0, which represents operations as designed, with one trainer responsible for training X1 to X2 learners.
- Regime -1, which represents under-utilization of capacity, with fewer than X1 learners for one trainer.

Regimes 1 and -1 are undesirable, as the former degrades service quality while the latter is not financially sustainable.

$$Regime = \begin{cases} 1 & \text{if } Learners \geqslant K_{Trainers} \text{ or } Learners \geqslant K_{Facility} \\ -1 & \text{if } Learners \leqslant X1 \\ 0 & \text{if } X1 \leqslant Learners \leqslant X2 \end{cases} \tag{5}$$

USE CASE

We now provide a specific example of applying the model above. The chosen example concerns the influence of the educational system's efficiency on maximizing the flow of learners to complete a full course in a French vocational training establishment, and its consequences for trainers' critical resources.

The *"Compagnons du Devoir"* vocational training system is based on the principle of traveling to learn a trade by gaining experience in different companies. After initial training, young company employees and work-study students travel from one training center to another, in France and abroad, to acquire the knowledge and skills specific to the trade they are learning. The training, accommodation, and catering facilities have been managed by associations of former beneficiaries of the scheme since the Middle Ages, forming a collaborative network for multiple trades where knowledge, know-how and interpersonal skills are exchanged.

As a highly skilled workforce, spread across the network of companies for which they work, the people trained by the *"compagnonnage"* system contribute to the restoration of infrastructures damaged by natural disasters. This takes the form of underpinning and reinforcement work, waterproofing when roofs are damaged, or repair work on industrial sites, depending on the type of damage. Each of these capabilities, taken together, will enable a building, an infrastructure component, or a piece of equipment to return to a functional state, thus ensuring the resilience of society as a whole.

"Compagnonnage" needs to design and ensure its resilience through the transmission of specific skills that can only be learned through experience. Over the last ten years, the number of apprentices joining the AOCDTF has doubled from 1,500 to 3,000. The training cycle lasts between 5 and 7 years, and the vast majority of trainers come from the end of the training course. Once they have become *"Compagnons"*, and have benefited from free,

Table 2. Assumptions on simulation initialization parameters

Parameters types	Definition	Values
$\tau_{Learners}$	the duration of the training	6 years
τ_{Senior}	the minimum time before leaving the system as a Graduate	3 years
$\tau_{Graduated}$	the minimum time before becoming a Trainer	1 year
$\tau_{Trainer}$	the minimum time before leaving the teaching system	2 years
$\Phi^t_{Trainer}$	the trainer transformation rate: graduated that will become trainer	50%
$\beta^t_{Learner}$	the learner abandonment rate (attrition rate)	35%
$\beta_{Graduated}$	the graduated abandonment rate, before becoming Senior or Trainer	10%
$\alpha^t_{Learner}$	the input of learners each year	417 per year
$\alpha^t_{Trainer}$	the input of external trainer each year	result of balance

high-level training, highly-skilled young workers commit to staying with the organization for 3 years, passing on the knowledge they have acquired to younger workers. This is the "Duty" that the *"Compagnons"* will carry out by becoming a trainer, or by continuing to travel and supervise the system of evening classes enabling workers to be promoted. 90% of young journeymen respect this commitment. They are the critical resource for the smooth running of the system. This 3-year transition time between graduation and leaving the system is taken into account and could be compared in the academic world to the situation of a doctoral student who for 3 years after obtaining a degree will continue to progress in his or her field of expertise until leaving the training process.

By simulating the training process for the *"Compagnons"* organization via the mathematical model derived above, we aim to answer the question: *"What happens when the attrition rate changes"* or *"What is the impact of improving the success rate?"* Studying the behavior of these variations will enable us to estimate the limits and regime changes facing the organization. We are basing our model on real data from *"Compagnons"* carpenters, in order to verify its conformity with reality.

We therefore experiment with the mathematical model, as implemented in Python, varying the attrition rate ($\beta^t_{Learner}$) to characterize the associated changes in system behavior. The initial conditions are given in Table 2.

RESULTS AND DISCUSSION

As shown in Table 3, we can see the variation in the annual attrition rate (AR), which accumulated over 6 years gives us the success rate (SR). We can see from the simulation results that, with an annual attrition rate of 35% of the workforce, if we start with 417 learners, after 6 years we still have 31 carpenters who have become *"Compagnons"*. Considering that 50% have the capacity to be trainers, this means a potential number of between 15 and 16, which is not enough to cover the annual requirement of 17 carpenter trainers. When we increase the attrition rate to 36%, we find that the number of *"Compagnons"* carpenters falls by a further 3 units, worsening the trainer deficit to -3 positions and -4.5 positions for 37% attrition. Similarly, by varying the attrition rate downwards, the simulation results show that the trainer balance becomes positive again at 33%, even releasing a potential of 2 additional trainers for each percentage point gained.

As listed in Table 1, in our future work we will be testing a range of risk factors specific to maintaining human capital, i.e. the lack of trainers, the attrition rate, and demographics, by testing extreme scenarios. By simulating the flow of people, we can also test the economic impact on the organization, as well as the budgetary limits of these fluctuations. As far as systems engineering is concerned, this flow simulation system contributes to planning, continuous improvement, updating strategy, and projecting the organization's administration into the future. All of these risk factors, which represent a range of real-world issues, give us confidence that organizations addressing these risks will become more resilient.

By testing these risk factors, we will accumulate information that will enable us to better understand the behavior of knowledge and know-how transmission systems. This will help us to more fully answer a variety of research questions, such as those posed at the beginning of the Model section, and thus help us to more effectively characterize the sustainability, limits, speeds and crisis sensitivity of human skill and knowledge chains.

Table 3. Carpenters simulation results

AR: Attrition Rate; SR: Success Rate; TR: Trainer Requirements; TB: Trainers Balance

AR/year(%)	SR after 6 years(%)	*Compagnons*	Variation	Annual TR	TB
37	6.0	25	−6	17	−4.5
36	6.7	28	−3	17	−3.0
35	7.4	31	0	17	−1.5
34	7.9	33	2	17	−0.5
33	8.6	36	5	17	1.0
32	9.6	40	9	17	3.0
31	10.6	44	13	17	5.0
30	11.5	48	17	17	7.0
29	12.5	52	21	17	9.0

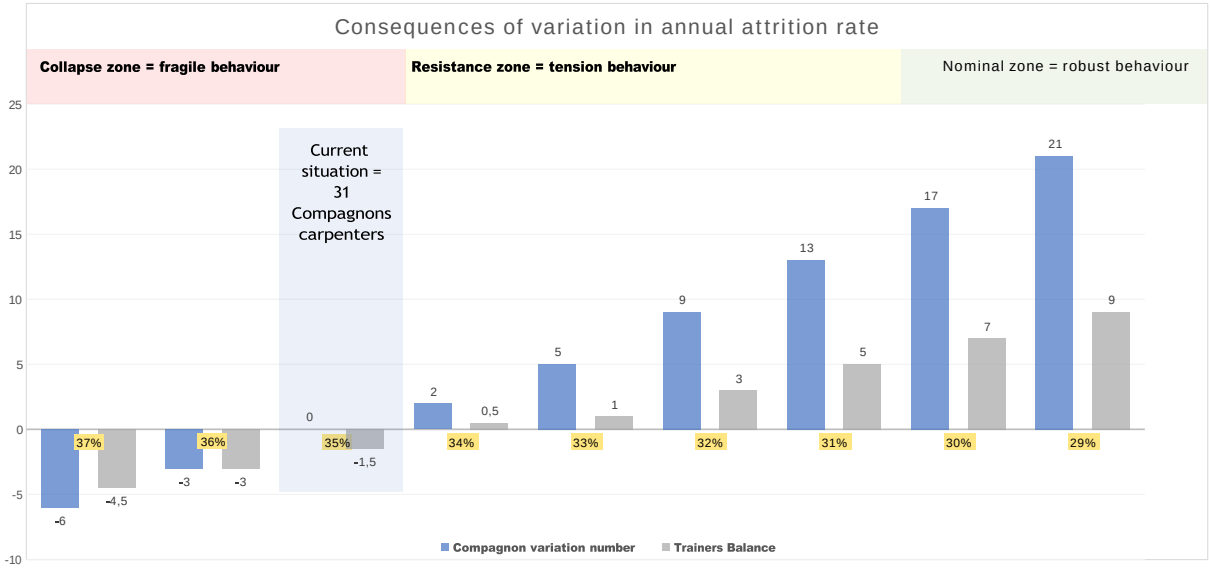


Figure 2. Attrition rate variation

ANALYSIS

Figure 2, which shows the variation in the percentage of attrition, enables us to analyze the preliminary results. The x-axis shows the variation in the attrition rate between 37% and 29% of the workforce, while the y-axis shows the workforce balance. The current benchmark of 35% is unsustainable, as it generates a chronic shortage of trainers. We have decided to position the simulations showing a trainer deficit in the collapse zone, which reveals the system's fragile behavior. We have arbitrarily chosen to place the scenarios between 34% and 31% in the resistance zone, which reveals that the system is under stress. We consider that in the case of *"Compagnons"*, from 30% attrition per year, the system remains in the nominal zone, suggesting robustness with constant behavior.

MODEL LIMITATIONS

This is a first version of our vision of the behavior of a learner flow in a know-how chain perspective, and like any model under construction, it does not claim to predict an indisputable result in an exact and precise manner. However, it is already useful for understanding the principles of circulation of flows of people transmitting knowledge, and can be used to estimate whether the system is in tension or not.

Attrition rates are parameterized linearly, whereas they tend to follow an irregular slope, with a sudden loss at the end of the 2nd year approaching 75%. This will be taken into account in the next iteration of model improvement, but in our case has no impact on the final number of trainers. This will help us to assessing the potential stock of learners, which is of interest for estimating the sizing of future infrastructures and the demand for trainers.

At this stage, we have not yet modelled regime changes. The various experiments we are going to simulate should enable us to define a high and low threshold for assessing the resilience of the transmission chain as practiced by the

"Compagnons". Empirically, it is generally accepted that an apprentice section is viable from 12 learners upwards, and that it becomes difficult to supervise more than 20 learners without degrading the quality of supervision, notably for safety reasons. We will integrate the consequences of these regime variations by giving ourselves the possibility of increasing the attrition rate as soon as the regime moves into overcapacity, and to parameterize a section collapse rule in the case of understaffing, starting from limits to be defined during our future simulations.

The model was initially implemented in Python, and a range of different outcomes were generated deterministically by varying the attrition rate and the number of incoming learners. Future extensions include further incorporating stochastic behavior in order to better characterize the range of possible outcomes.

CONCLUSION AND FUTURE RESEARCH DIRECTIONS

As mentioned above, the *"Compagnons organization"* is currently experiencing a growth crisis of varying severity, depending on the trade. For Carpenters, the average annual need for trainers is 17. If we consider that one out of every two *"Compagnons"* has the capacity to become a trainer, we can immediately see a chronic deficit of 1.5 trainer positions, which accumulates over the 2-year cycle and is confirmed by the interviews conducted with *"Compagnons"*. By varying the loss rate downwards in 1% steps, we can see the effects of making the balance positive from 33% loss, by releasing a sufficient number of journeymen with the capacity to be trainers, and thus covering the need.

These results show how the organization is behaving, and make it more straightforward for managers to become aware of the situation by providing them with the tools they need. Simulating learner flow makes it easier to understand the issues involved in maintaining the ability to transmit know-how or knowledge. This insight highlights the importance of questioning the *"Compagnons"* training system, so that it pays greater attention to the success rate of learners, which is one of the keys to renewing this human chain of know-how. It also indicates that this type of model could be helpful in other long-term training contexts where shortages are already identified, such as in healthcare services or civil security.

With regard to the model, we will continue our efforts to improve it, taking into account the fact that the attrition rate is not linear, so as to be able to forecast the number of learners, grads, and trainers per year with greater reliability. We will also implement the possibility of simulating a regime change to vary success rates by an additional factor. This will bring us closer to real-life behavior, where we can observe a deterioration in success rates when staffing levels are insufficient, or when a training course eventually disappears due to low enrollment. Ultimately, many factors need to be identified and combined in order to fully answer our research questions.

It is, of course, also important to recognize that many different types of crises can disrupt education systems over extended periods of time, such as armed conflicts resulting in conscription of young people (collapse of incoming numbers), demographic crises (fewer births and consequent drop in incoming numbers), financial crises (drying up of apprenticeship funding), vocational crises (poor image of manual trades), reputational crises (trades perceived as degrading for society following a news item resulting in a bad reputation), or public health crises (confinement preventing professional practice). Furthermore, although we have focused here on such slow-motion disasters, sudden-onset disasters such as floods, wildfires, major storms, or terrorist attacks can also impact the proper functioning of an education system in a number of different ways. The impacts of such events could also be incorporated into the simulation model through their effects on training times and abandonment rates; however, to fully explore this more traditional disaster management context it will help to develop a more detailed flow model that incorporates more of the logistics of the training process.

The differentiation of the organization's behavior into three types - fragile, resistant or robust - still needs to be qualified and tested in our next iterations of the model. Once we have gathered the available data to align as closely as possible with reality, we will be able to test the behavior of different scenarios. We will implement an experimental design by varying key parameters, thereby gaining the ability to make recommendations.

In order to evaluate the resilience of the knowledge chain, we will test our model to define its resistance thresholds. A design of experiment will therefore be conducted to observe the system's absorption capacity according to scenarios of input flow and attrition rate variability. We will also observe its recovery time under different hypotheses. Finally, we will assess the potential for adaptation in order to prescribe recommendations for improving the overall resilience of the chain and avoiding disruptions that have consequences in terms of loss of know-how and societal capacity to cope with natural disasters.

There are 36 *"Compagnons"* trades, each with a different learner flow behavior and resilience level. Once the model has been tested, we'll be able to diagnose this behavior, make appropriate recommendations, and assist decision-makers by simulating the potential impact of their wishes. Once all 36 scenarios have been aggregated, we

will also be able to identify a trend in the resilience behavior of the *"Compagnons"* organization as a whole, first of all to trainer shortages, but ultimately to a wide variety of different types of disruption. This, in turn, will support the organization's ongoing efforts to continue training new members and it will allow them to continue providing the necessary expertise to make our critical infrastructures more resilient.

ACKNOWLEDGEMENTS

Thanks to Virginia Tech, Maggie Webb, Dr Marie Paretti, and Dr Mathieu Besanc¸on for support and advice, and to the *"Compagnons"* from the AOCDTF for funding this research work.

REFERENCES

Bruneau, M., & Reinhorn, A. (2006). Overview of the Resilience Concept. *Proceedings of the 8th U.S. National Conference on Earthquake Engineering*.

Burner, T. (2018). Why is educational change so difficult and how can we make it more effective? *Forskning og forandring*, *1*(1), 122–134. https://doi.org/10.23865/fof.v1.1081

Dehdashti, A., Fatemi, F., Jannati, M., Asadi, F., & Kangarloo, M. B. (2020). Applying health, safety, and environmental risk assessment at academic settings. *BMC Public Health*, *20*(1), 1328. https://doi.org/10.1186/s12889-020-09419-5

Draus, P. J. (2009). Substance Abuse and Slow-Motion Disasters: The Case of Detroit. *The Sociological Quarterly*, *50*(2), 360–382. https://doi.org/10.1111/j.1533-8525.2009.01144.x

Intergovernmental Panel On Climate Change. (2023, July). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009157896

Jacobson, M. J., Levin, J. A., & Kapur, M. (2019). Education as a Complex System: Conceptual and Methodological Implications. *Educational Researcher*, *48*(2), 112–119. https://doi.org/10.3102/0013189x19826958

Jorro, A. (2007). L'e´valuation ge´ne´ratrice de de´veloppement professionnel ? L'Harmattan.

Meerow, S., Newell, J. P., & Stults, M. (2016). Defining urban resilience: A review. *Landscape and Urban Planning*, *147*, 38–49. https://doi.org/10.1016/j.landurbplan.2015.11.011

Michel, A. (2000). Évaluer pour piloter. *Revue internationale d'e´ducation de Se`vres*, (26), 19–29. https://doi.org/10.4000/ries.2521

Nadoveza Jelic, O., University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia, & Gardijan Kedzo, M. (2018). Efficiency vs effectiveness: An analysis of tertiary education across Europe (University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia, Ed.). *Public Sector Economics*, *42*(4), 381–414. https://doi.org/10.3326/pse.42.4.2

National Research Council. (2012). *Disaster Resilience: A National Imperative*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13457

Nielsen, L., & Faber, M. H. (2021). Impacts of sustainability and resilience research on risk governance, management and education. *Sustainable and Resilient Infrastructure*, *6*(6), 339–384. https://doi.org/10.1080/23789689.2019.1611056

OECD. (2023, September). *Education at a Glance 2023 Sources, Methodologies and Technical Notes*. https://doi.org/10.1787/d7f76adc-en

Pierrehumbert, R. T. (2006). Climate Change: A Catastrophe in Slow Motion. *Chicago Journal of International Law*, *6*(2).

Saikia, R., & Bezborah, D. P. (2014). Quality and Relevance in Higher Education. *5*(2).

Sandborn, P. A., & Prabhakar, V. J. (2015). The Forecasting and Impact of the Loss of Critical Human Skills Necessary for Supporting Legacy Systems. *IEEE Transactions on Engineering Management*, *62*(3), 361–371. https://doi.org/10.1109/TEM.2015.2438820

Valinejad, J., Mili, L., & Van Der Wal, C. N. (2022). Multi-Agent Based Stochastic Dynamical Model to Measure Community Resilience. *Journal of Social Computing*, *3*(3), 262–286. https://doi.org/10.23919/JSC.2022.0008

Waller, R., Lemoine, P., Mense, E., Garretson, C., & Richardson, M. (2019). Global Higher Education in a VUCA World: Concerns and Projections. *Journal of Education and Development*, *3*, 73. https://doi.org/10.20849/jed.v3i2.613

Barriers against data sharing for Disaster Risk Reduction (DRR) at borders

Anouck Adrot
Université Paris-Dauphine, Université PSL, CNRS, DRM, 75016 PARIS, France
Anouck.adrot@dauphine.psl.eu

Lola Brunet
Université Paris-Dauphine, Université PSL, CNRS, DRM, 75016 PARIS, France
Lola.brunet@dauphine.psl.eu

Henri Isaac
Université Paris-Dauphine, Université PSL, CNRS, DRM, 75016 PARIS, France
Henri.isaac@dauphine.psl.eu

ABSTRACT

Data sharing is a major driver to disaster Risk Reduction (DRR), which is essential to cross-border regions' resilience. However, sharing data for DRR at borders remains challenging and a limited spectrum of regions have been significantly progressing on this path. While the literature recommends that cross-border regions and organizations settle data ecosystems, studies also evidence that ecosystems face difficulties in developing sustainable rules for data sharing. A better understanding of the impediments of data sharing in cross-border DRR ecosystems is therefore necessary and remains missing in the literature. This research, based on a qualitative and collaborative design, is a work in progress focusing on the Italian-French border. Based on 24 qualitative interviews and archive analysis, we have identified four major challenges that organizations need to overcome to support data sharing for DRR at borders. First, cross-border DRR actors struggle harmonizing their perspectives over their common environment, whether it is environmental or da ta-related. Second, they face persistent elements of fragmentation, particularly reinforced at frontiers. Third, these organizations meet the challenge of building a sustainable bond of trust between one another about data. Fourth, they lack data maturity, thus impeding good use of data within the ecosystem. Our findings contribute to a more situated understanding of data sharing at borders and ca ll for more consideration of the border settings from practitioners that participate in data sharing.

Keywords

Disaster Risk Reduction (DRR), cross-border, data sharing, natural risk, ecosystems.

INTRODUCTION

On the 13th of August 2023, heavy rains in the Savoy region in France generated flash muds in Bardonecchia, in Italy. Also, the rise of temperatures generated major landslides in the Alps, for instance on the 27th of August at La Praz, next to Modane, in the Maurienne valley. These events disrupted railways and highways that support transportation, workers mobility and economic exchanges between Italy and France. The main direct railway between France and Italy for now is and will remain close until the end of 2024. These examples illustrate how vulnerable to disasters cross-border regions can be.

Cross-border regions comprise natural separations such as rivers and mountains. They tend to get marginalized from the center of countries both economically and politically (Şlusarciuc, 2015). However, they frequently host critical infrastructures such as highways, dams, industrial and energy plants. They can even attract exploratory investment, such as in the case of the economic corridor that has been developing between Pakistan and China (Lim, 2019). For these reasons, the capacity of organizations at borders to prevent and address disasters has become a major concern for local and national authorities but also international organizations (Abad et al., 2018). Therefore, all these actors have been working at supporting cross-border DRR, which consists of identifying and

preventing risks that can generate disasters (Palliyaguru et al., 2014). DRR also encompasses collective response to these disasters.

Data can drive DRR cooperation across borders (Kanbara & Shaw, 2021). First, as a digital resource, it is easy to replicate and share. Second, it calls for collaboration because understanding and preventing disasters cannot rely on a single source of data (Petrenj et al., 2013), as illustrated in the case of Bardonecchia . Indeed, the heavy ra ins that struck the Italian village had started on the other side of the border, in France.

What is data sharing? Data sharingdoes not boil down to sporadic sending or reception of data sets. As explained by Jussen et al. (2023), data sharing refers to the collective organizing of data access within a specific ecosystem. To that extent data sharing is sociotechnical and does not restrict to technological interoperability (Liverani et al., 2018). Many barriers can prevent data sharing within ecosystems (Gelhaar et al., 2021), going from the lack of understanding of the value of data to fears regarding data security. In the context of cross-border DRR, data sharing has been already documented but studies on the topic from a technical stance lack (e.g. Toro et al., 2019), leaving the topic of data sharing barriers insufficiently explored (Petrenj et al., 2013).

Given the foregoing, this work addresses the following question: " *What is the nature of barriers against cross-border data sharing for DRR?"*

This work aims at addressing this question by relying on a qualitative research design. We have considered the case of the Italian-French border, which has been struggling with the increasing frequency of disasters triggered by landsides and floods in the Alps. While Italy and France have been separately working on elaborating data bases on natural risks and vulnerabilities, there is no official data sharing system between the two countries. This research is still in progress and requires further enrichment. So far, data has been collecting through 24 interviews, archives and meetings with experts and practitioners. We have been following grounded theory principles (Glaser & Strauss, 1967; Suddaby, 2006). Some preliminary findings have been emerging and outline four major barriers against data sharing at borders: lack of a shared vision of DRR and data, fragmentation, lack of trust and lack of data maturity. We detail each barrier as well as the interdependencies between the barriers. We then discuss the novelty of these findings, the future avenues to improve the research as well as their practical implications.

THEORETICAL BACKGROUND

Disasters and DRR at borders

Borders are particularly exposed to disasters and critical situations. Most borders are shaped by topological elements prone to disruptive events. For instance, water zones – like the Rhin between Germany and France – or the Tanganyika Lake between Democratic Republic of Congo and Tanzania- are subject to floods. As another example, mountains – like the Alps between Austria and Italy or the Himalayas between China and India- are subject to landslides and earthquakes. Cross-border regions have been increasingly vital for economic integration, in all the continents (Brunet-Jailly, 2022). Paradoxically they also become more vulnerable because they host critical infrastructures, such as nuclear plants, energy infrastructures, dam s and transportation channels.

For this reason, cross-border regions – defined here as territories at the jointure between distinct countries separated by a topological or administrative border – need to become resilient. DRR can drive the strengthening of this resilience while achieving sustainable development by preventing new and reducing existing and new disaster risks (Palliyaguru et al., 2014).

There has been discussion of the key concepts of DRR (Palliyaguru et al., 2014). Disaster risk corresponds to the possibility of harmful consequences resulting from the interactions between hazards and vulnerable conditions (UN-ISDR, 2009). In this context, vulnerability covers both exposure to hazards and capacity to address them. In other terms, vulnerability is the combination of, on the one hand, the degree of exposure of a system (societal, productive, etc.) and on the other hand its capacity to prepare for and respond to the hazard (UN-ISDR, 2009). It is important to note that hazards are not purely natural. Indeed, floods, fires and landslides can be favored – if not triggered - by urbanistic design and human intervention.

Based on the review provided by Palliyaguru et al., DRR can be approached as a set of four interdependent strategies, namely i) eliminating (avoiding the risk), ii) mitigating (which means reducing the impact of hazards, as well as their frequency, intensity and scale), iii) preparing (developing resilience and response capacities and plans) and iv) advocacy (favoring good practices through influence). Mitigating disaster related risks thus depends on decades-long projects that involve international and national planning (McConnell & Drennan, 2006). Also, advocacy is a long-term stake of DRR because it consists of lobbying institutions, organizations and communities.

Furthermore, the management of crises and disasters at borders requires collaboration between a large spectrum of organizations: companies, administration, agencies, communities and authorities, from multiple countries and

sectors (Ansell et al., 2010; Ansell & Boin, 2009). In line with this view, DRR at borders involves a large spectrum of stakeholders, from local to national and international actors. At a local level, municipalities endorse the responsibility of citizens but also of access to commodities including water, transportation, etc. Communities are involved to promote well-being and sustainability of economies, which also depend on infrastructures. At a national level, states are responsible for the protection of their population and the mobilization of their emergency actors. At an international level, transborder organizations can encourage - legally or not - the use of good practices. Increasingly, DRR actors have been aiming at integrating the diversity of processes and approaches to disaster risks, but without significant success (Abad et al., 2018).

Data Sharing for DRR at Borders

A critical resource to manage disasters is reliable data (UNDRR, 2015). Organizations involved in disaster response hardly dispose of the data that can address all their informational needs (van den Homberg et al., 2018). For this reason, organizations need to work on data preparedness (van den Homberg et al., 2018) by sharing data before a disaster strikes (Sarker et al., 2020). Data used for DRR can be very diverse, ranging from natural and industrial risks to demographic data. However, challenges remain regarding timeliness, completeness and precision of data.

Growing use of data for DRR illustrates the search for digital maturity worldwide, supported by the United Nations (for more detail check the UNDRR report 2022). While some countries in the world have been working on centralized disaster data systems, other countries have been promoting collaboration to share data. These countries have put efforts in developing data ecosystems (van Esch, 2021) and open data (Kanbara & Shaw, 2021).

In the scholar sphere, data sharing has been investigated as a socio-technical phenomenon observable in a large spectrum of sectors (see for instance Hazell et al., 2023 and Ure et al., 2009). This means that data sharing does not restrict to the settling of infrastructures and programs that can make data sets interoperable. Rather, data sharing is also a human practice that involves multiple organizations (especially for DRR) and requires willingness and efforts to harmonize data and make it accessible to others (Migliorini et al., 2019). However, there has been a persistent knowledge gap regarding the use of multiple data sets (for instance provided, consumed and exchanged in a data ecosystem) in the programming of resilience (Migliorini et al., 2019).

Therefore, data sharing represents a major component of resilience. At European borders, projects such as ESPRESSO or GIOCONDA (Toro et al., 2019) have outlined the benefit from data sharing across the border. Also, the European commission has been supporting the rise of data ecosystems by producing data frameworks to help organizations address the diversity of rules and standards about data (European Commission. Joint Research Centre., 2021). Beyond Europe, the World Economic Forum (WEF) has supported the sharing of data between Turkÿe and Greece to produce an early warning system about fires. In Asia, such at the border shaped by the Mekong river, cross-boundary organizations were created to promote data sharing between countries (Thu & Wehn, 2016).

The mutualization of data can better help grasp the nature and the amplitude of a risk. Once the crisis strikes, data sharing can help coordinate actions (Liverani et al., 2018). Even when tensions arise between countries, organizations still persevere to collect and share data (Thu & Wehn, 2016), which also reveals the importance of data sharing. However, sharing data for DRR at borders remains easier said than done. The literature has reported many difficulties. First, data sharing, as a socio-technical phenomenon, is complex (Liverani et al., 2018). It relies on a set of technologies whose interoperability needs to be designed (Jussen et al., 2023; Toro et al., 2019). Going further, the literature has documented the diversity of issues inherent to data sharing in ecosystems (Jussen et al., 2023)

Despite growing interest in data sharing in the context of cross-border DRR, there is a lack of comprehensive understanding of the challenges inherent to data sharing at borders. This calls for more investigation of the nature of phenomena that can prevent data sharing. In the following section we detail the methodology that we have been following to address the research question.

METHODOLOGY

Research Design

As a reminder, this work addresses the question of the nature of the barriers against data sharing for DRR across borders. To do so, we have been conducting a qualitative research design, which focuses on the case of the Italian-French cross-border regions. We detail in this section our choices and the challenges that we met to collect data. We also present our analysis practices and tools. In a nutshell, this study embeds into a wider collaboration with

a Non Governmental Organization (NGO) to conduct an action research project. This study is also inclusive and has been involving master students.

Case Study: the Italian-French Cross-Border Region

To address the research question, we have been conducting a case study focusing on the Italian -French border region, which comprises multiple valleys (such as the Suza, Vésubie and Roya).

The Italian- French cross-border regions have been impacted by natural hazards several times. The two countries are separated by the Alps and share a common Mediterranean coast. The latest disasters were the most impactful. For instance, in October 2020, Alex storm impacted the cross-border territory between Liguria (Italy) and Alpes-Maritime (France). Alex destroyed many infrastructures (including cemeteries and roads) as well as houses. The damages from Alex are still being addressed. In August 2023, a major railway connection between Lyon and Torino was damaged by a landslide. As a result, transportation of goods had to be deviated towards Vintimiglia in the South of Liguria. Last summer, massive waves of mud damaged roads, cars and buildings in Bardonechia. These waves resulted from particularly heavy raining on the French side.

To understand barriers against data sharing in this case study, we first identified the actors involved in DRR in the region. Data collection from both a literature review and some interviews (see below) led us to list over 70 entities, of which 31 are Italian, 37 are French, 5 are international actors and 5 are European . Interestingly, the diversity of the DRR ecosystem implies the possibility for organizations from various sectors to share data (hea lth, transports and energies, security of people and goods, etc.). The ecosystem also comprises organizations with complementary interests in data (from consuming to producing and communicating). However, few interactions between organizations at the border about DRR involve data sharing. Organizations might tend to cooperate when they are linked hierarchically or functioning. But sharing a same territory does not incent organizations to share data.

Data Collection

We have been collecting data from both experts and individuals from organizations that participate into DRR at the Italian-French border. During our interviews, we have been relying on a dictionary theme and an interviewer guide that have been iteratively refined. In the interviews, we first ask the interviewee to present one's work and contribution to DRR and resilience. Some interviewees were specialized in territory cohesion, some others in natural risks and others in responding to disasters. We have been collecting data on the various steps of DRR. In parallel, we also relied on various archives (such as articles and reports on the use of data for DRR at borders). Even though cross-border data sharing for DRR is not extensively documented in the literature yet, many organizations (such as the European Commission in Europe) as well as communities have been organizing events to promote collaborative usage of data for DRR at borders. Table 1 presents the interviews that we have already conducted, it highlights the diversity of interviewees, in terms of occupation and profiles. Roughly, we did our best to interview experts who have knowledge and expertise of cross-border cooperation, natural risks and DRR, as well as subjects who had some practical experience of cooperation for DRR and data sharing. Finally, we interviewed people from various nationalities, including Italian, French, but also German and Canadian.

Approaching the topics of data and data sharing has been a challenge in our study. While experts can have a precise idea of the data needed for DRR (in particular at borders), they do not always have an exhaustive view on the data practically used by organizations. Similarly, some operationals could not tell us which data they use and whether these data have been shared. Our difficulty collecting precise and exhaustive data about data remains unsurprising. As documented, data is not a trivial asset for orga nizations. Most of the time, its access remains confidential. To address the difficulty of interviewees expressing themselves about data, we multiplied interviews. We also made a parallel search in 2023 to find data sets that were accessible on the web (through portals and open data sources). We then identified the organizations that provided the data and contacted them. We also relied on snowball sampling to identify organizations that shared data with the interviewees. By doing so, we have progressively improved the richness and relevance of the interviews and we almost reached semantic saturation. Our next step is to refine our coding through cross-coding and conduct the interviews necessary to reach semantic and theoretical saturation.

Despite our focus on the Italian-French frontier, we have been expanding data collection on multiple borders in Europe and beyond. For instance, we have been considering the cases of the Greece-Turkÿe, as well as the frontier formed by Mekong River between Thailand, Lao, Cambodia and Vietnam. This has helped us identify whether the barriers identified in our data related to the Italian-French border specifically or manifested in other cross-border regions as well. This practice has helped us refine our findings and supports the reliability of it.

Table 1. List of the interviewees

N°	Interview ID	Occupation	Duration	Date
1	16_01_Exp_01	Scholar in Climatic Geopolitics	44:28	16/01/23
2	02_03_Op_01	Civil protection commandant	01:13:15	02/03/23
3	28_02_Exp_04	Scholar in Humanitarian studies	01:21:09	28/02/23
4	27_01_Exp_02	Humanitarian project manager	59:56	27/01/23
5	27_02_Exp_03	Chief Officer of a project on resilience and critical infrastructures	01:04:10	27/02/23
6	23_06_Op_02	Emergency manager	37:38	23/06/23
7	10_07_Op_05	Emergency manager	01:27:32	10/07/23
8	27_06_Exp_07	Scholar in geography, risks and data	44:38	27/06/23
9	03_07_Op_04	Firefighter Captain	42:54	03/07/23
10	29_06_Exp_08	Scholar in geography, risks and data	01:03:20	29/06/23
11	16_08_Op_06	Emergency Manager	01:03:00	16/08/23
12	23_06_Op_03	Prefectural decision maker	51:03	23/06/23
13	29_12_Exp_09	Former manager of cross-border projects	50:00	29/12/23
14	04_01_Exp_10	Director of natural risks prevention unit	01:32:38	04/01/24
15	18_01_Exp_11	Territorial Cohesion advisor	49:09	18/01/24
16	26_01_Op_07	Chief Officer of a service devoted to risks and territories	01:26:13	26/01/24
17	02_02_Op_08	Emergency Manager	01:17:43	02/02/24
18	05_02_Exp_12	Natural risk expert	01:16:12	05/02/24
19	05_02_Exp_13	Scholar in cross-border resilience	01:29:15	05/02/24
20	06_02_Exp_14	Chief officer of service devoted to natural risks	01:06:05	06/02/24
21	06_02_Exp_15	Scholar on cross-border cooperation	01:31:46	06/02/24
22	13_03_Exp_06	Scholar in Climate Change Response Strategies	01:00:13	13/03/23
23	10_03_Exp_05	Scholar in International Humanitarian Law	59:21	10/03/23
24	14_02_Exp_16	Director of non-profit organization	45:00	14/02/24

Data Analysis

So far, we have been relying on qualitative coding of the interviews on MaxQDA to identify themes and let them emerge in an inductive modeling. We are still processing the analysis of the collected data and it seems that, as said, we have almost reached semantic saturation.

Our data analysis has been following the principles of grounded theory. First, based on literature reviewing and focus groups, we identified a main concern, which corresponds to the practical issue that the subjects of the study intend to address (Glaser and Strauss 1967). Second, we systematically compared data across interviews. By doing so, our findings emerged from data. We also based our reasoning on abduction by putting into perspective the various sources of data. For instance, based on literature reviewing and focus groups, we elaborated a theme dictionary that we iteratively refined when completing interviews.

We found several ways to support the reliability of our coding. First, we relied on dialogical reasoning to challenge each other's views to refine our findings. Second, in April 2024, we will present these findings during a workshop

gathering both some experts and some operational actors. This will be an opportunity to discuss and refine our findings.

FINDINGS

The findings of this research remain preliminary and require refinement. Though, we present in this section four interrelated barriers against data sharing for DRR across borders. We have been inferring them from our data analysis. First, multiple legal, institutional and political legacies co-exist at a border. In addition, countries work on different standards and methods that do not necessarily meet. We label this phenomenon fragmentation. Fragmentation results in poor knowledge and awareness of other organizations' needs and data usage at the border. It thus contributes to the lack of mutual trust especially related to data use. Also, organizations are not equally acculturated to data in relation to DRR. This implies the lack of mutual understanding of data usage between organizations and thus the lack of trust regarding data. The lack of data maturity among organizations also undermines the building of a shared vision of data for DRR. Finally, both lack of mutual trust and shared vision of DRR and data prevents organization from addressing fragmentation.

These four challenges are depicted in Figure 1. Each barrier is symbolized by a box. The arrows represent how one barrier fuels another. In the findings section we detail each barrier but also their interdependencies.

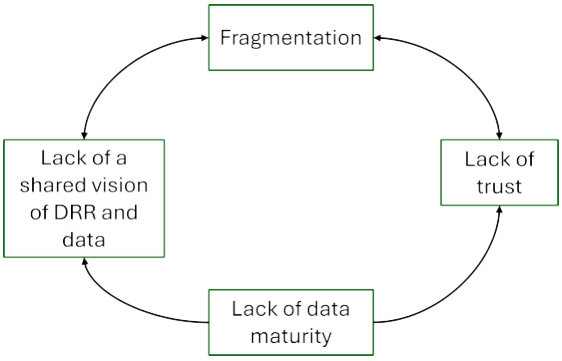


Figure 1. Interdependent barriers against data sharing at the Italian-French border

Fragmentation at Borders

In this work, fragmentation refers to the divides, differences and inconsistencies of views and practices between multiple stakeholders. Fragmentation manifests in the lack of cooperation but also technical inconsistencies (organizations using different standards), organizational and cultural divides. Fragmentation can also stem from different languages. Borders are naturally fragmented territories because they comprise natural or artificial separations. Even on the same side of a border, divides and rivalries exist between organizations.

In a fragmented ecosystem, organizations might be interested in the same topic but never interact with each other. An expert explains how fragmentation, by preventing communication between DRR actors, impedes the development of a shared knowledge on the risks and the methods to address them :

"France is specific in that responsibilities are spread between various organizations. In France, these organizations do not communicate with each other. This means that organizations do not share their knowledge on risks inherent to a specific territory. This also means that no organization has a complete understanding of the risks. This electricity provider, for instance, focuses on the risks that are the most important to its functioning and relies on its own methodologies. It is the same for train transportation companies. The State manages national roads, and the issue is the same. Highways? The same. And so on. No exchange". [26_01_Op_07].

This quotation highlights that each organization deals with the risks it could be affected by, whereas their concern should also extend to the risks that can affect other organizations and the other side of the border. Fragmentation represents a vicious circle. The less organizations are aware of the multiplicity of risks (and methods), the more they focus on their own risks and methods. And the less they become aware of the systemicity of risks on a cross-border territory. Because organizations hardly communicate on risks, they hardly develop an overview of risks at the border. This means that in the long run, fragmentation undermines the emergence of a shared vision related to DRR. Fragmentation is also a problem in terms of data sharing because it prevents the emergence of a common representation of data, both from a technical perspective (formats, metadata that is necessary, means to store the data) and from an organizational/social perspective (how to use data, methods to transmit the data).

Lack of a Shared Vision of DRR and Resilience at Borders

Cross-border territories are often characterized by the presence of topological obstacles, whether these obstacles are embodied by a river, a mountain or a desert for instance. Organizations from the diverse sides of the border might not identify the same risks and/or not prioritize risks in the same order. This is the case at the Italian-French border, where Italy has developed stronger vigilance towards volcanoes' activity and earthquakes than France.

Simultaneously, Italian and French plans might struggle to coexist in a unique framework. To nuance our view, though, a shared vision of cross-border DRR is emerging. The two countries have recently experienced cross-border disasters (such as Alex storm and 2023 landslides). They both agree on the need to tackle these risks. However, this vision of cross-border DRR turns out fragile. Primarily, DRR is still perceived as a local issue and cross-border territories remain traditionally approached as two separated territories rather than one single liminal space. An expert of the Italian-French border explained that the Alps remain a line and a separation between two territories rather than a single shared territory. In addition, DRR actors suffer from a lack of clear overview on how the Italian-French cross-border area will practically tackle disaster risks in the future. The building of a shared vision implies the need to redefine the cross-border territory.

This prevents DRR actors from making choices about the future avenues to support data sharing. Frequently, frontiers are primarily represented as lines of separation. This is how they appear on maps and official documents for instance. Historically approached as a separation, the Italian French border was not an object that required attention. Therefore, collaboration for cross-border DRR requires the renewal of the vision of the border.

Also, beyond these topological matters, the lack of vision concerns what data could be helpful for and what usage should be settled in priority. Even though DRR actors share common concerns across the border (for instance, how to reduce vulnerabilities, how to evacuate people and how to allocate resources), a shared vision of what data could help achieve remains missing and prevents the development of a shared data culture.

In other terms, the absence of a shared vision impedes data sharing. Indeed, most interviewees outlined that the rise of a cross-border strategy for DRR should be inclusive and involve proactive collaboration. In particular, DRR can be co-constructed by actors that share diverse types of expertise. As explained by an expert frequently involved in resilience projects, co-constructing solutions to address disaster related risks appears very beneficial. Co-construction of solutions implies repeated interactions that provide the opportunity for DRR actors to understand each other's roles and values well:

"[Organizations] rely on us because we have some expertise and some data to bring. We bring technical expertise to them. When we deliver this expertise, it is clear that we increasingly co-construct the final product or service [with these organizations]. This is the way it is. Co-constructing things is a way for us to be sure that we will address the demands [from organizations]. It is a way to well understand what [organizations] need and to be sure that we well address needs and to understand how our services can be used for what action afterwards". [26_01_Op_07]

Co-construction offers DRR actors not only an avenue to know each other, but also an opportunity to understand how each actor contributes to a collective effort and develops an inclusive strategy for DRR. The absence of shared visions, by impeding such co-construction, restrains the development of trust that is necessary for data sharing. Because of siloed visions, actors tend to interact only with their closest counterparts, which can also aggravate fragmentation.

Lack of Trust

Our empirical findings suggest that cross-border cooperation requires dynamic and continuous engagement from every side of the border. Developing resilience is impossible without full participation of citizens and communities. Cooperation at borders requires trust and working on mutual trust, as suggested by both these two interviewees:

"If they need us, they have to trust us and if we want to work to help some other partners (...), we need to believe in it (...). And if we think that there is no trust, I think we should rather put the matter on the table and tackle it or clarify our intervention". [26_01_Op_07]

"Fundamentally, this is what we look for, trust, as it is essential to cooperation, which itself is essential in order to solve these problems". [18_01_Exp_11]

Interviewees suggest that trust is not accumulative but rather dynamic. This means that organizations need to keep testing their ties, for instance through collaboration on data . From this perspective data can represent a relevant resource and driving force to cooperation on DRR. However, settling and maintaining cooperative projects over time across a border is demanding for organizations, particularly because of politics and the potential legacy of conflicts at frontiers. Conflicts can also question organizations' participation into data sharing projects without

notice.

To illustrate this point, our archive analysis reveals that Greece and Turkey face common natural hazards such as earthquakes, extreme droughts, volcanic eruptions, mudslides, flash floods, forest fires and potential tsunamis in the Mediterranean Sea. Despite the need for bilateral cooperation to mitigate these risks, the relationship between the two countries is characterized by a strong historical antagonism that hinders risk and disaster reduction. Past cooperation agreements on DRR were endangered by the continuing political tensions. In the future, additional communication breakdown between Greece and Turkey would likely compromise DRR cooperation and data sharing.

Moreover, the interviewees highlight the fact that data has a contradictory effect on trust. On the one hand, data represents a valuable resource to collaborate on risks because, as mentioned earlier, one single organization (or country) cannot produce on its own all the data necessary to predict disasters. On the other hand, data can be a lso used to produce manipulative narratives. As explained by an interviewee:

"So data is not data. Data is information about something that happened or might happen and it's controlling narratives. So what I have seen and what the literature also says is that people is more afraid about the narrative than sharing the data. So is the data saying that I'm good at it? It's going to say that I'm under control, it's going to say that I'm not under control in the case, it's saying that I'm capable to do something. So it's about the narrative, the data more than the technicality of how to share data. The other part that is more technical, but it's not just technicalities is about data for itself. Although it's a narrative, it doesn't talk for itself and that's what a lot of actors are really afraid of sharing data because if you don't know the context in which the data was collected, for which purposes, you can quite easily misuse it or misunderstand it and it's almost super difficult to explain all the context. (…) [Data's] power is narrative (…) it doesn't tell the whole story". [28_02_Exp_04]

In sum, data sharing calls for trust and is essentially collaborative. However, trust about data remains fragile at frontiers. In addition to persistent sources of defiance related to politics and conflict's' legacy, data can be perceived as a major threat. Therefore, DRR actors need to overcome this fragility, which complicates cooperation and data sharing. From this perspective lack of trust represents a major barrier against data sharing.

Data Maturity Within the Ecosystem

Data maturity is of utmost importance to promote data sharing in cross-border regions. At the Italian-French border, the relationship between organizations and the concept of data is imperfect and uneven. DRR actors and organizations are not always properly trained in data tools and processes. Therefore, these organizations lack adequate tools to experiment and collaborate about data. As a result, some participants might collect data, but without any perspective on its future use and sharing. Unawareness of the importance of data sharingmight lead to situations where data has been collected yet not shared before a disaster. As explained by an expert of border resilience:

"*I think it's a question of knowledge and information, of being well-aware which type of data, how many different data, what means strategic data or operational data… I think, at the end of the day, what we have is a lack of knowledge, of how to handle data*". [14_02_Exp16]

Eventually, data collected without any perspective on its usage eventually becomes obsolete and irrelevant. Organizations need to update data on a regular basis to keep it relevant. At the cross-border scale, this means that interoperability between data sources and meta data should be constantly ensured. Even when updated, data needs to be kept in the data collector's "language".

Interdependencies Between the Barriers

Our empirical data suggests that the four barriers outlined here interact and self-feed themselves. These interactions are more detailed in the following lines.

First, the fragmentation and the absence of a shared vision of DRR and data are mutually dependent. Fragmentation renders organizations ignorant about one another. Overcoming divides and rivalries - possibly inherited from past conflicts at the border - requires doubling efforts to identify synergies between data sources. In parallel, as outlined by the interviewees, the absence of a shared vision implies that organizations get closer to their closest partners. This attitude, in its turn, fuels fragmentation and the setting up of "bubbles" of actors rather than a diversified and transversal ecosystem.

Second, fragmentation prevents organizations from building trust which, in turn, aggravates fragmentation Cross-borders areas are characterized by their plurality of equipment and processes. Making it interoperable might not

be doable, at least in the brief term, thus preventing actors from understanding each other's methods and mistrusting it. As a result, each actor tends to rely on his own practices, fueling the existence of silo es and fragmentation.

Third, data immaturity impacts both the existence of shared visions and of trust. Of course, not understanding what data is and how it can be used makes it impossible for a common vision of the potential use of data to emerge. At the same time, as they perceive data as a "niche" matter far from them, actors under-experiment with it, yet it is precisely experimentation that makes it possible to build trust. Two other interactions implying data immaturity can be imagined, but still need some more evidence. First, we might think that the lack of shared visions deters organizations from acculturating to data: as they don't have a common goal they wish to reach, they are not encouraged to support the costs of training. Second, lack of trust in other participants' might persuade organizations to keep distant from their data and more generally from data "literacy".In short, solutions to one barrier shouldn't be considered independently from those to the other ones. Even when a shared vision is reached, it must be accompanied by a behavioral adaptation (tools, language, etc.) that will lighten fragmentation. What's more, these technical/practical issues are not sufficient: the relational dimension of trust seems to be fundamental. Eventually, these "ease" in relationships between actors should reflect in an ease in the relation ship with data, made possible by a process of data acculturation.

DISCUSSION

Data sharing has been identified by institutions and scholars as a major driving force to DRR. Data sharing at borders does not seem effective yet at multiple borders. The nature of the barriers against data sharing for DRR at borders remains partially known. For this reason, this research addresses the following research question: " *What are the barriers against data sharing for DRR at borders?"*

Our empirical findings outline four major barriers, namely i) visions of data and territories that do not meet, ii) elements of fragmentation particularly heightened at border, iii) the paradoxical building of trust about data at borders and iv) insufficient development of data acculturation. In this section, we first detail the expected contribution of the research. Then we infer some theoretical insights from our findings. Finally , we briefly present some practical contributions from the first findings.

Expected Contribution From the Research and Avenues for Future Development

The research is still in progress. In the future months, we intend to pursue the analysis of the collected data to propose a comprehensive model of the barriers against data sharing for DRR at borders. Ideally, the comprehensive model will help induce recommendations for organizations at the Italian-French border to strengthen data sharing and cooperation for DRR. As mentioned herein, this research is part of a wider project that aims at the development of a sustainable data ecosystem at the Italian -French border. The results from the research will be used as a scientific basis to design a governance plan for the ecosystem.

Theoretical Insights: DRR Ecosystems and Data Sharing

Our findings outline the importance of fragmentation between the Italian-French border. This finding echoes previous research on borders (Kratke, 1999). However, this work goes a little further by exploring the implications of cross-border inconsistencies on data sharing. So far, the literature has outlined that data sharing at borders confronts a lack of interoperability, divergent vocabularies, norms and equipment (Liverani et al., 2018). Those are the elements we referred to as "fragmentation" in this pa per. However, the literature has more rarely explored some deeper root causes of operational issues, such as the absence of common perspectives on territories and data. This research aims at providing a comprehensive view on these root causes. By doing so, we hope to better help organizations address invisible yet powerful barriers against data sharing at borders.

Also, the literature has explained how organizations can gather in ecosystems to organize and govern data sharing for DRR (van Esch, 2021). By doing so, they can develop many to many – rather than one to many - connections. Not only they enlarge the spectrum of usable data. Also, they can mutualize resources to host data and invest into infrastructures. Our empirical findings outline fragmentation, lack of trust and lack of data maturity as serious obstacles to data sharing, which could even block any initiative to govern data sharing. Developing ecosystems at borders require the tackling of issues that are not inherent to the ecosystems itself, but rather related to the cross-border settings they have been developing in.

Practical Implications

Our research also outlines the importance of shared efforts to build trust between actors: they cannot be designed in a top-down fashion. More precisely, our preliminary results suggest the need for a bottom -up, inclusive and practical leadership on data sharing. As empirical findings suggest, DRR actors across the border need to dismantle issues that can undermine cooperation on a very regular basis.

Also, our research suggests the need for cross-border DRR actors to overcome fragmentation and the lack of trust through the design of rules to share data. Future development of this research will focus on how to integrate fragmentation and lack of trust into a governance plan. Practically, this means the need to identify which rules can aggravate fragmentation of views on data and which rules can, on the contrary, contribute to a unified vision of data sharing. Similarly, there is the need to identify which rules can undermine or foster DRR actors' trust towards each other about data.

CONCLUSION

This research aims at better understanding the nature of the barriers against data sharing for DRR at borders. Based on a qualitative research design, it has been exploring the case of Italian and French data sharing for DRR at the border. This research is still in progress but outlines four persistent barriers that need to be tackled: i) the lack of shared visions of yet common environments, ii) the persistence of elements of fragmentation, iii) the fragility of trust in relation to data and iv) lack of familiarity with data. While still preliminary, these findings suggest the need for DRR actors to increase their knowledge of data usage but also its barriers at borders. Also, it suggests the need to incorporate such knowledge in the design and maintenance of data governance between two or more countries.

While the literature advocates the creation of ecosystems to ease data sharing (European Commission. Joint Research Centre., 2021), we alert on the limitations of "by-the-book" solutions to do so. Adoption of frameworks and tools without considering the features of the cross-border region that hosts the ecosystem can be misleading. On the contrary, at borders, DRR actors need to collaboratively construct a specific vision of DRR and data. This implies the need to overcome siloed visions, fragmentation, lack of trust and underdeveloped capabilities in relation to data. In its future development, this research aims at providing a more comprehensive understanding of the barriers against data sharing for DRR at borders.

ACKNOWLEDGMENTS

This research has been financially supported by the European Commission (EC) and is part of the ECHO Red Roses project (101101221). We also thank the following master students who participated into the research project: Capucine Baud-Berthier, Karima Bennani, Juliette Briens, Romane de la Séglière, Juline Fayard, Marie Hou, Violetta Pagani, Pierre Michaud.

REFERENCES

Abad, J., Booth, L., Marx, S., Ettinger, S., & Gérard, F. (2018). Comparison of national strategies in France, Germany and Switzerland for DRR and cross-border crisis management. *Procedia Engineering*, *212*, 879–886. https://doi.org/10.1016/j.proeng.2018.01.113

Ansell, C., & Boin, A. (2009). *Managing Transboundary Crises: Requirements for a Dynamic Response System.*

Ansell, C., Boin, A., & Keller, A. (2010). Managing transboundary crises: Identifying the building blocks of an effective response system. *Journal of Contingencies and Crisis Management*, *18*(4), 195–207.

Brunet-Jailly, E. (2022). Cross-border cooperation: A global overview. *Alternatives*, *47*(1), 3–17.

Data and digital maturity for disaster risk reduction: Informing the next generation of disaster loss and damage databases (p. 84). (2022). United Nations Office for Disaster Risk Reduction United Nations Development Programme (UNDP). http://www.undrr.org/quick/75391

European Commission. Joint Research Centre. (2021). *Establishment of sustainable data ecosystems: Recommendations for the evolution of spatial data infrastructures.* Publications Office. https://data.europa.eu/doi/10.2760/04462

Gelhaar, J., Groß, T., & Otto, B. (2021). *A taxonomy for data ecosystems.*

Glaser, B., & Strauss, A. (1967). Grounded theory: The discovery of grounded theory. *Sociology the Journal of the British Sociological Association*, *12*(1), 27–49.

Hazell, P., Novitzky, P., & van den Oord, S. (2023). Socio-technical system analysis of responsible data sharing in water systems as critical infrastructure. *Frontiers in Big Data*, *5*, 1057155.

Jussen, I., Schweihoff, J., Dahms, V., Möller, F., & Otto, B. (2023). Data sharing fundamentals: Characteristics and definition. *Proceedings of the 56th Hawaii International Conference on System Sciences (HICSS). Maui, Hawaii, USA.*

Kanbara, S., & Shaw, R. (2021). Disaster Risk Reduction Regime in Japan: An Analysis in the Perspective of Open Data, Open Governance. *Sustainability*, *14*(1), 19. https://doi.org/10.3390/su14010019

Kratke, S. (1999). Regional integration or fragmentation? The German—Polish border region in a new Europe. *Regional Studies*, *33*(7), 631–641. Arts Premium Collection; ProQuest One Business.

Lim, A. C.-H. (2019). The moving border of the China -Pakistan economic corridor. *Geopolitics*, *24*(2), 487–502.

Liverani, M., Teng, S., Le, M. S., & Coker, R. (2018). Sharing public health data and information across borders: Lessons from Southeast Asia. *Globalization and Health*, *14*(1), 94. https://doi.org/10.1186/s12992-018-0415-0

McConnell, A., & Drennan, L. (2006). Mission Impossible? Planning and Preparing for Crisis1. *Journal of Contingencies and Crisis Management*, *14*(2), 59–70. https://doi.org/10.1111/j.1468-5973.2006.00482.x

Migliorini, M., Hagen, J. S., Mihaljević, J., Mysiak, J., Rossi, J.-L., Siegmund, A., Meliksetian, K., & Guha Sapir, D. (2019). Data interoperability for disaster risk reduction in Europe. *Disaster Prevention and Management: An International Journal*, *28*(6), 804–816. https://doi.org/10.1108/DPM-09-2019-0291

Palliyaguru, R., Amaratunga, D., & Baldry, D. (2014). Constructing a holistic approach to disaster risk reduction: The significance of focusing on vulnerability reduction. *Disasters*, *38*(1), 45–61. https://doi.org/10.1111/disa.12031

Petrenj, B., Lettieri, E., & Trucco, P. (2013). Information sharing and collaboration for critical infrastructure resilience–a comprehensive review on barriers and emerging capabilities. *International Journal of Critical Infrastructures*, *9*(4), 304–329.

Sarker, M. N. I., Peng, Y., Yiran, C., & Shouse, R. C. (2020). Disaster resilience through big data: Way to environmental sustainability. In *INTERNATIONAL JOURNAL OF DISASTER RISK REDUCTION* (Vol. 51). ELSEVIER. https://doi.org/10.1016/j.ijdrr.2020.101769

Şlusarciuc, M. (2015). The economic potential of crossborder areas. Opportunities and threats. *Procedia Economics and Finance*, *32*, 801–808.

Suddaby, R. (2006). From the Editors: What Grounded Theory is Not. *Academy of Management Journal*, *49*(4), 633–642. https://doi.org/10.5465/amj.2006.22083020

Thu, H. N., & Wehn, U. (2016). Data sharing in international transboundary contexts: The Vietnamese perspective on data sharing in the Lower Mekong Basin. *Journal of Hydrology*, *536*, 351–364. https://doi.org/10.1016/j.jhydrol.2016.02.035

Toro, J. F., Carrion, D., Albertella, A., & Brovelli, M. A. (2019). CROSS-BORDER OPEN DATA SHARING: GIOCONDA PROJECT. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, *XLII-4/W14*, 233–238. https://doi.org/10.5194/isprs-archives-XLII-4-W14-233-2019

UNDRR. (2015). *Sendai Framework for Disaster Risk Reduction 2015 -2030*. https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030

United Nations International Strategy for Disaster Risk Reduction (UN-ISDR) (Ed.). (2009). *Terminology on Disaster Risk Reduction*.

Ure, J., Procter, R., Lin, Y., Hartswood, M., Anderson, S., Lloyd, S., Wardlaw, J., Gonzalez-Velez, H., & Ho, K. (2009). The development of data infrastructures for ehealth: A socio -technical perspective. *Journal of the Association for Information Systems*, *10*(5), 3.

van den Homberg, M., Monné, R., & Spruit, M. (2018). Bridging the information gap of disaster responders by optimizing data selection using cost and quality. *Computers & Geosciences*, *120*, 60–72. https://doi.org/10.1016/j.cageo.2018.06.002

van Esch, S. (2021). *Looking Beyond the Data: An Assessment of the Emerging Data Ecosystem of Nepal's Flood Early Warning Systems*. 12.

End-user perspective on the results and impact of European projects: a case of failure to innovate.

Maaike van de Vorst
APB Campus Vesta
Maaike.vandevorst@campusvesta.be

Kenny Meesters
Tilburg University
K.J.M.G.meesters@tilburguniversity.edu

Bert Brugghemans
Brandweerzone Antwerpen
Bert.brugghemans@bza.be

ABSTRACT

First responder organisations (FROs) participating in European research projects often seem dissatisfied with their participation. This is because they are unable to implement the project results. This article looks at this phenomenon through the lens of ambidextrous organisations. It examines the different factors contributing to failure to innovate for FROs in EU projects by using interviews. The main results show that the structure of European projects is a major contributor to failure to innovate of the project results in FROs. Nonetheless, the structure of the FROs themselves can also provide hinderances.

Keywords

End-users, first responder organisation, failure to innovate, ambidexterity, project management.

INTRODUCTION

For the period of 2021-2027 the European Commission has invested over 1,8 trillion in European research projects (European commission, Directorate-General for budget, 2021a). Examples include the Seventh Framework, HORIZON 2020, HORIZON EUROPE, ISF, EUCPM, ESA and many other initiatives. One of the underlying principles of these frameworks is to not only invest in fundamental research, but also foster innovation and accelerate the adoption of new academic insights into practical applications. As the European Commission outlines: *"bringing about tangible results and positive change in the lives of citizens and other beneficiaries is, and will remain, a key priority for the Commission in all the interventions supported by the EU bud get."* (European commission, Directorate-General for budget, 2021a , P4.)

This extends to specific programs aimed at increasing the resilience of the European Union in the field of emergency response and crisis management, among others in the Disaster Resilient Society program (DRS). Especially in these programs the connection between research and valorisation of the results to support first responders' organisation is strongly advocated. This is, among others, embodied by the program's requirement to include first responder organisation in the consortia that are submitting their project proposals under the call.

The question about the impact of these programs has been asked before (European Commission, Directorate-General for Research and innovation 2018). This paper examines specifically if the impact is applicable to the daily practise of first responder organisations, and if participating in EU projects is of added value for first responder organisations (FROs). This way, it examines if there is a disconnect between the intended outcomes of the projects and the actual impact they have in FROs.

Problem Description

There may be multiple reasons why such a discrepancy occurs. First, the EU projects and programs may have (too) strong a focus on innovation and research with a lower direct applicability to the objectives and operations of first responder organisations. However, this is not only addressed in the call for proposals, but even a strong requirement. For example, a certain technology readiness level (TRL) in Innovative Actions is expected to ensure a higher and quicker adoption rate (Rösner, 2023). Second, innovation and research programs may have insufficient connection to 'practice'. This is however also addressed in the programs by ensuring participation of First Responder Organisations in the consortiums as well as specific dissemination plans addressing primary end users (i.e. FRO's) throughout the project. Furthermore, participating FROs may lack resources to fully engage in these projects and/or adopt their results (Meijers et al., 2006). Again, measures such as increased funding rates are aimed at addressing these potential shortcomings. Most importantly: it is also not a shortcoming due to a lack of need, motivations or drive to participate in these projects. The need to innovate, adapt and gain new knowledge is well understood by FROs. It is evidenced by the following quote from one of the respondents in this research:

"[innovation] is essential. The world in which we work evolves quickly. As an organisation we need to stay relevant in this world we work in (…) it is not the question if we should, it is evident we should."

As the quote above from one of the respondents demonstrates, implementation of research and innovation (RI) is necessary for first responder organisations working in our fast-paced society. This article examines implementation of such project outcomes in FROs. Specifically, by focusing on the organisational capabilities and capacities to successfully leverage the potential of innovation projects in general, and EU-funded research programs aimed at first responder organisations in particular.

Research Questions

As mentioned in the introduction, there seems to be a failure to innovate European results at FROs. Thus, one objective of this article is finding out why that is the case. Therefore, the first research question (RQ) asks:

1. Why is there seemingly a failure to innovate of European Project results at first responder organisations?

The conceptual model in figure 2 shows two main relations that are of interest. The one between EU projects and project results, and the one between project results and implementation in FROs. If there is failure to innovate, there must be a breakdown in one of these relationships which interrupts the innovation stages. Therefore, the analytical interviews are used to examine where this breakdown occurs, why it occurs and what factors cause it to occur. The second research question goes one step further and asks what can be done about that breakdown.

2. Which action can first responder organisations take before, during and after European projects to prevent failure to innovate?

To answer RQ2, a handbook was developed for FROs to use before, during and after European projects which includes worksheets that can be used. Thus, the handbook should help FROs feel better prepared to work on EU project, but it should also make them aware of the pitfalls that may be common to FROs.

THEORETICAL BACKGROUND

Ambidexterity and First Responder Organisations

When considering the role first responder organisations play in society, a strange dichotomy presents itself. On the one hand, FROs need to react quickly to a crisis, while maintaining coverage of their district to respond to other incidents. On the other hand, FROs need to be prepared for the future, and the crises that come with that. In other words, FROs need to be ambidextrous. Preparing for the future also means being able to implement new RI in the organisation. Ambidexterity originated in the literature as a concept meant for companies. It is the ability to both maximise current profits and sales but are also being flexible and adaptable enough to respond to the future by implementing RI (O Reilly & Tushman, 2004; Birkinshaw & Gibson, 2005). Nowadays scholars also apply it to public organisations and non-for-profit organisations. For example, Narayanan and Altay (2021) use ambidexterity to describe humanitarian organisations like food banks and the salvation army. Others use it to describe to the functioning of the army during war time versus during peace time (Soeters 2008). Thus, applying it to FROs has precedent.

Another aspect of FROs is that they are often public in nature. Thus, the theoretical background takes into account literature around RI implementation for ambidextrous organisations, but also for public organisations. This information is fused with the results of the exploratory interviews, creating a conceptual model considering the unique characterises of a FROs as both a public organisation and an ambidextrous organisation.

FROs in European Projects

In this article "European projects" is used for all projects that are completely or partially funded by Europe (European commission, Directorate-General for budget, 2021a, 2021b). This article distinguishes 3 phases for European projects, based on the exploratory interviews and the literature. Firstly, the proposal phase, secondly the execution phase and thirdly the result phase.

- Proposal phase: during this phase the proposal is being written, and the consortium is created. Important steps in this phase are: 1) finding relevant partners, 2) dividing the budget, 3) dividing the tasks, 4) describing the expected impact. The proposal is then evaluated (European commission, 2018, 2021c, 2021d, 2022a, 2022b; Cunningham, O'Reilly, Hooper, Nepelski & Van Roy, 2020).
- Execution phase: during this phase the Grant Agreement needs to be executed. End -users (the role of participating FROs) often enact the following tasks:
    - Defining gaps in the work field
    - Formulating end user requirements
    - Scenario building
    - Validation of results
    - Dissemination of results (European commission, 2022c; 2022d)

  One of the respondents mentioned the following*: "FROs work on the practical side. We organise field trials, we test with the work field if something works or not. We start after the research has been done."*
- Result phase: in this phase the official project period has ended, and so had the financing. During this phase, the results should be implemented in either the participating organisations or the member states (European Commission, European Innovation Council and SMEs Executive Agency, Scherer, Weber & Alveen, 2022a).

According to Serwiak (2022), not all end users seem satisfied with the project results of EU projects. The following difficulties are often encountered:

1. Not enough resources to participate
2. EU projects are not part of the core tasks
3. Too little experience with EU projects
4. It takes too long for results to be usable
5. Shifting priorities of end users during a project
6. Bad experiences in previous projects (Serwiak, 2022)

Reasons to participate in European projects are similar for FROs as the reasons to participate in RI implementation. They want to improve the efficiency or effectiveness, involve citizens or corporations, increase safety and equality, or increase their own legitimacy (De Vries, Bekkers & Tummers 2014; Rip, 2016).

Factors Influencing Implementation

The process of innovation has multiple stages. In this article, 4 stages are identified. These are the initial stage, the development stage, the adoption stage and finally the implementation stage (Somech & Drach -Zahavy, 2013; De Vries, Bekkers and Tummers, 2014; Allen et al., 2017; Singh, Akbani & Dhir 2020). These stages overlap with the phases of European projects as seen in figure 1.

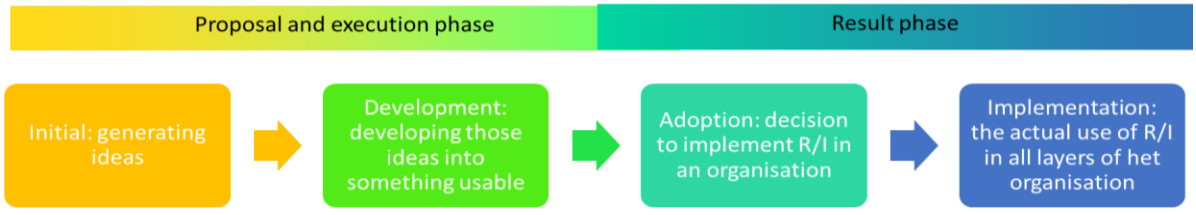


Figure 1. Overlap between innovation stages and EU project phases

The factors that influence research and innovation implementation in FROs can be divided into four categories. These are the characteristics of the project results, individual factors, organisational factors and environmental factors. The table below show the factors that can influence implementation (simplified). A "+" refers to a positive relation with implementation, while a "–" refers to a negative relation. The factors in bold have been emphasized because of their importance for RI implementation. There are a few factors that have a +/- sign, in this case there is either debate on the positive or negative relation, or it could work both positively or negatively depending on the other factors. This table does not show possible interactions between the factors.

Table 1. Factors influencing implementation

	Implementation (literature)	**Implementation (exploratory interviews) for European projects**
Characteristics RI (Meijers et al, 2006; De Vries, Bekkers, Tummers, 2014; Alexander, 2015; Owen, 2018; Brooks, Grugulis & Cook, 2021; MED1stMR, 2023 respondents)	- Abstraction + Complexity + Concrete solutions + Congruency with current way of working + Ease of use **++ Added value** - Costs + Testability + Access	+ Adaptability +/- Low TRL level
Individual (Mulgan & Albury, 2003; Starkweather, & Kardong-Edgren, 2008; Somech, & Drach-Zahavy, 2013; De Vries, Bekkers, & Tummers, 2014; Alexander, 2015; Hero, Lindfors, & Taatila, 2017; Owen, 2018; Moussa, McMurray & Muenjohn, 2018; Cunningham et al., 2020; Singh, Akbani & Dhir, 2020; respondents)	+Creativity +Communication skills +Flexibility	**++Combination of project management skills and operational experience as FR** +Communication skills +Flexibility
Organisational (Young, Charns & Shortell, 2001; Mulgan & Albury, 2003; Meijers et al., 2006; Edmond, Megivern, Williams, Rochman & Howard, 2006; Büschgens, Bausch, & Balkin, 2013; De Vries, Bekkers, Tummers, 2014; Jacobs et al., 2015; Mosadeghrad, 2015; Fu, Liu & Liao, 2018; Moussa, McMurray & Muenjohn, 2018; Owen, 2018; González Fernández, Kubus & Mascareñas Pérez-Iñigo, 2019; Clausen, Demircioglu, & Alsos, 2020; Cunningham et al., 2020; Singh, Akbani & Dhir, 2020 Stelzl, Röglinger & Wyrtki, 2020; Brooks, Grugulis & Cook, 2021; Koskinen, 2021; Akıncı, Alpkan, Yıldız & Karacay, 2022; Smara & Bogatyreva 2023; respondents)	+Active unlearning +reward for experimentation or risk taking +Demonstration and trainings +Diversity - Memories of the past - Hierarchy +Open and closed leader +Leader: specific and concrete vision of implementation. ++**Resources for R/I** +Encourage development of employees - Complexity **++Time for R/I** ++**Safe working climate**	+ Composition of the consortium **++Extra resources for R/I from EC** +Shared responsibility for project results in the consortium +/- Vision of project coordinator +/- Vision of participating consortium members
Environment (Young, Charn & Shortell, 2001; Strejcek & Theil, 2003; Mulgan & Albury, 2003; De Vries, Bekkers, & Tummers, 2014; Mosadeghrad, 2015; European Commission, Directorate-General for Research and innovation, 2018; Clausen, Demircioglu & Alsos, 2020; Brooks, Grugulis & Cook, 2021, respondents)	+participation in isomorphic networks +/- Economic pressure +/- Political pressure - changes in pressure	- European environment +/- societal relevance

Conceptual Framework

Based on the above literature and the exploratory interviews, a conceptual model can be created where the presence of factors is integrated with figure 1. This can be seen in figure 2. The figure shows that the EU project has influence on the project results that are achieved. The factors present in FROs influence that relationship via their participation in the project. Those project results then need to be implemented during the result phase. In this phase, the factors present in FROs influence that as well.

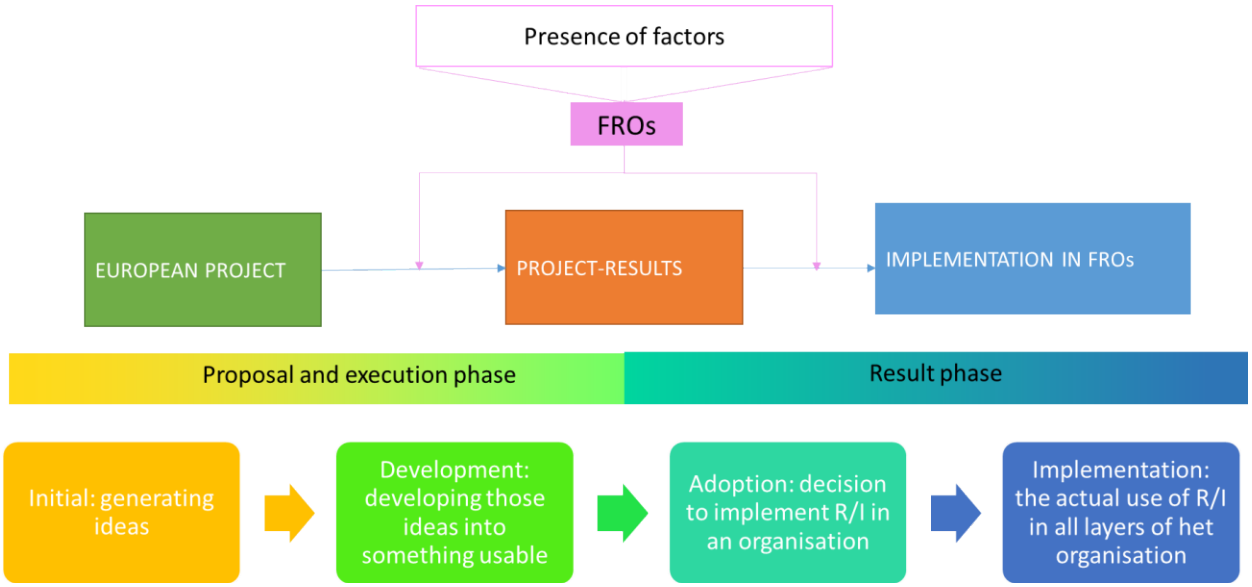


Figure 2. Conceptual model of FRO implementation of EU project results (own design)

RESEARCH APPROACH

Figure 3 shows the methodological approach in this article. Firstly, literature was found about implementation in public organisations and in ambidextrous organisations. On top of that, a round of five exploratory interviews was conducted to complete the theoretical framework. That framework exists out of a conceptual model and the first version of a handbook for FROs[1]. This model and the handbook were then evaluated in a second round of six analytical interviews. Based on those interviews, the conceptual model and the handbook were adjusted to their current version.

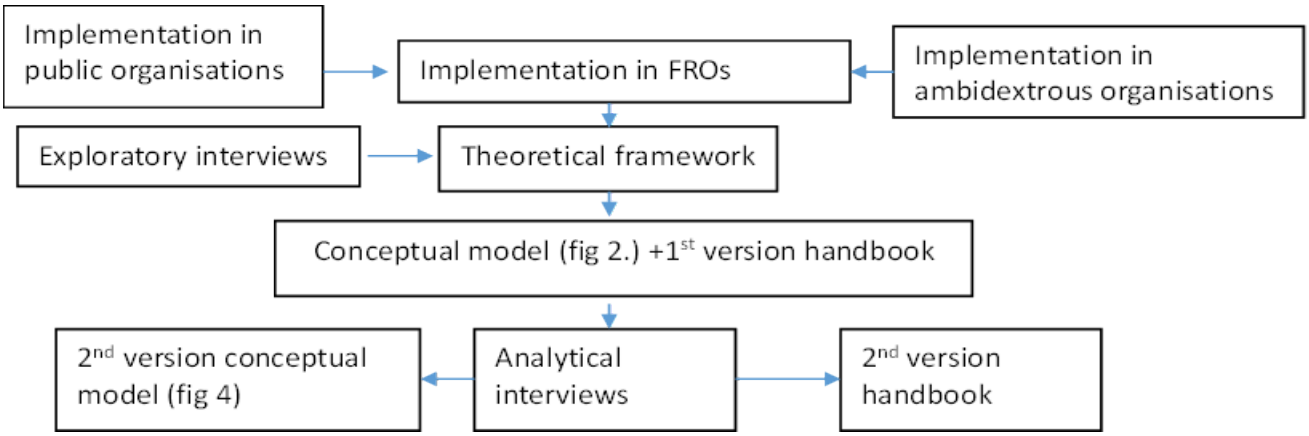


Figure 3. Methodological approach (Own design)

Data Collection

To choose what literature to use in this paper, we set up an excel table that tracked a ll found literature. It was then checked for inclusion of: RI implementation elements, if the end user was a FRO and if the paper addressed European projects. The articles in this paper include at least 2/3 elements. On top of that, priority was given to meta-analysis to get an overview of the different factors that influence implementation.

There were two rounds of interviews done for this research. The first round was the exploratory interviews. The

[1]Since this was only a concept version, it is not included in this article.

goal of these interviews was to gain a better understanding of the relationships in the conceptual model. For both interviews, we used the snowball method. On top of that, we tried to create a good spread of respondents, including both managerial roles as well as operational roles. The exploratory interviews were open conversations, belonging to the category of non-structured interviews. The shortest exploratory interview was 36 minutes and the longest was 1 hour and 12 minutes. The interviews were conducted in Dutch. 4 of 5 respondents were from the Flemish part in Belgium, and one was from the Netherlands. Of those exploratory interviews, 1 worked for the fire services and was a trainer in a training centre, 1 worked for the police department of a training centre, 1 worked for the national police, 1 for the medical services as well as being a trainer in a training centre and 1 as an advisor EU projects for a training centre. The exploratory interviews were transcribed via Amberscript. They were coded via the inductive method. Text fragments of interest were chosen and pasted into excel. Every new concept was then assigned a code that best represented the text fragment. This was done to all text fragments of the exploratory interviews. Because many codes were identified, some were linked to each other in themes. these themes were derived from the literature review done before.

For the second round of interviews, the analytical interviews, 6 respondents were interviewed. All 6 respondents were from the Flemish part of Belgium. the shortest interview was 1 hour and 2 minutes, while the longest was 2 hours and 16 minutes. Of these respondents 1 was from the fire services, 1 an advisor EU projects for a training centre and from the fire services, 2 worked for the police, 1 worked for a training centre, and 1 worked for a national crisis centre. All the respondents had experience with European projects, mostly with ISF (internal security fund) and HORIZON 2020 projects. In contrast to the exploratory interviews, the coding of the analytical interviews was done deductively. The codes and themes were based on those developed during the coding of the exploratory interviews. Nonetheless, new codes could still be added if necessary. This allowed for a structured but flexible framework through which to interpret the interviews. A semi-structured interview method was chosen for the analyticalinterviews because there were specific points that needed to be addressed. By using an interview protocol, we could make sure to discuss all those points. During the interviews there was room for new insights that could challenge the conceptual model made from the theoretical framework. Following an interview protocol but allowing for deviations made the semi-structure interview the best choice for this research.

Relevance and Validity

This research is exploratory in nature, more research is needed to explore the subject fully. The research is based on a total of eleven interviews, 5 exploratory and 6 analytical. The respondents of this research were found via the snowball method. The small number of respondents and the methods of finding respondents could lead to undesirable bias. It could be the case that the respondents are not representative for FROs working on EU projects, which could influence the validity of this research. On top of that, two authors work for a FROs that participate(d) in EU projects. This too could lead to unconscious bias, for example with confirmation bias.

RESULTS

For the analytical interviews, a total of 735 text fragments were encoded. Table 2 shows the top 10 codes[2] most associated with either the EU project structure or the FRO structure and table 3 shows the codes most used per project phase, mostly used for RQ2. Table 4 shows the top 3 codes associated with the themes. Interpretation of these results is found in the discussion.

Table 2. top 10 codes associated with project structure and FRO structure

EU projects structure codes top 10	**Amount**	**FRO structure codes top 10**	**Amount**
Profile project management	19	Vision on RI	37
Composition organisation	18	Management	28
Time	13	Resources	25
Administration	13	Added value	24
Management	12	Time	21
Resources	12	Innovation implementation	21
Profile expertise	11	Politics	20
Scale	9	Open attitude	20
Communication	8	Demo and trainings	12
Added value	7	Own initiative	12

Table 3. Most codes used per project phase

Project phase	**Top codes**
Call selection	Added value (5x) Concrete solutions (3x) Composition organisation (3x) Alternative(3x)
Proposal	Composition organisation (4x) Decide own objectives (3x) Profile project management (3x)
Execution	Administration (12x) Profile project management (10x) Composition organisation (9x)
Results	Innovation implementation (10x) Added value (9x) Concrete solution (9x)

[2] Translated from Dutch

Table 4. Most used codes per theme

Theme	Top 3 codes
EU project individual **Total: 42x**	Profile project management (16x) Profile expertise (11x) Communication (7x)
Individual **Total: 67x**	Own initiative (12x) Open attitude (12x) Profile expertise (8x) Ego (8x)
EU project organisation **Total: 123x**	Composition organisation (17) Time (13x) Administrative (13x)
Organisational **Total: 193x**	Vision on RI (33x) Leadership (28x) Resources (21x) Time (21x)
EU project environment **Total: 22x**	European rules and regulations (7x) Politics (4x) Society (4x)
Environment **Total: 79x**	Politics (19x) National networks (11x) Nationale rules and regulations (9x)
EU project characteristics **Total: 22x**	Added value (6x) Concrete solutions (5x) European collaboration (2x) Costs (2x)
Characteristics **Total: 23x**	Added value (13x) Abstraction (3x) Concrete solutions (2x) Ease of use (2x)
Use of project results **Total: 28x**	Innovation implementation (6x) Concrete solutions (5x) Added value (3x) Internal dissemination (3x)
Use of RI **Total: 17x**	Innovation implementation (5x) Research implementation (3x) Society (2x)
Advantages EU projects **Total: 30x**	Own network and contacts (5x) Resources (5x) Concrete solutions (4x)
Disadvantages EU projects **Total: 39x**	discouragement (7x) Abstraction (7x) Administration (7x)
Advantages RI **Total: 1x**	Reorganisation (1x)
Disadvantages RI **Total: 8x**	Reorganisation (4x) Efficiency (1x) TRL level (1x) Innovation implementation (1x) Changeability (1x)
Recommendations **Total: 41x**	Alternatives (11x) Composition organisation (6x) Decide own objectives (4x)

DISCUSSION/CONCLUSION

The Role of Characteristics

During the analytical interviews, it became clear that the characteristics play a crucial role in the implementation of results. The results needed to have added value and offer concrete solutions to relevant problems, but they often did not. Many respondents felt that the structure of EU projects was to blame for this. The following quote illustrates that: "*It often feels that end users in research project are an afterthought. We make a nice academic project, but we also need to involve first responders because of the commission. To tick all the boxes. So, we look for a few FROs that are not allowed to think in the front end about what we are investigating. We are going to research something that is academically relevant, the practical relevance does not matter.*"

Another respondent mentioned that you need to involve FROs to get the right characteristics. "*If you ask organisations to implement results that have been thought out only in the ivory tower, the results will be much worse than if they would have had actual good contact with our colleagues. Then you know, this has been developed by people like us. The moment it becomes abstract, it does not go well.*"

This section shows that if the results do not offer concrete solutions, added value or relevance, the implementation process stops after the EU project (RQ1). Thus, it is the creation of these characteristics during the project that is of importance. Furthermore, there are also hints that once a result is complete, the characteristics moderate the relationship between all the other factors (individual, organisational, environmental), and the implementation. De Vries, Bekers and Tummers (2014) mentioned that there is less interest in the influence of characteristics in the RI field for public organisations. This article shows that characteristics do play a major rule. Thus, characteristics especially should be considered a target for future RI research, and especially its possible moderating role.

Factors That Influence Characteristics in EU Projects

The next question that needs to be asked is which factors specifically are important for the creation of the characteristics. According to the respondents this is mostly the composition of the consortium, and the coordination of the project. Respondents mention (and are supported Cunningham, et al. (2020)), that often coordinators have a scientific or industry background, and that this influences their priorities. One of the respondents only participates in EU projects when they coordinate, while others mention that they do not have the means or experience to coordinate a project.

The composition of the consortium is another factor that influences the result. According to respondents it should be *"as big as needed and as small as possible"*. There is one respondent who prefers to work in consortia only with FROs, while the others believe that diverse consortia with different kinds of organisations are still the best to work with.

Respondents also mentioned that the European environment should do more to stimulate implementation of project results. For example, by representing the interest of end-users. Requiring their presence in a consortium is not enough. On top of that, it would be necessary to bridge the gap between the end of a project, when a result is (often) not ready to be implemented, and the actual time a result is implementable. Currently, the respondents feel that this is not being done sufficiently which harms the implementation. Furthermore, respondents feel that the administration that is needed to participate in EU projects is too time-consuming. One of them said: "*Of course you need to be accountable in EU projects. It is public money. But if you look at the archaic software in which everything needs to be reported and how you need to report the same thing three or four time, it takes many hours and many people working to do those things that have no use at all.*"

Apart from the interim evaluation of the different European project programs, there was little research on the specific factors that influence characteristics in EU projects. Thus, this can be a possible gap in the literature that should be studied more in depth.

Other Important Factors for FROs Participating in EU Projects

Other impact factors to consider for FROs that participate in EU projects are:

- Individual
  - Ego: the respondents mentioned that many FROs were proud of the job they do, and wanting to implement change could be read as criticising the current way of working.
  - Own initiative: this can be illustrated by the following quote: "*Research and innovation [in FROs] is completely non-committal and depends on individual interest and initiative"*.
  - Open attitude: the respondents remark that some first responders have a conservative attitude, while others have an "early adapter" mentality. Unlearning could be a way to convince FRs with a more conservative attitude, as discussed by Brooks, Grugulis & Cook (2021).
- Organisational
  - Time, resources, and a shortage of personnel are all hindrances to RI implementation in FROs. The importance of these practical factors is also underlined in the literature. The absence of these can completely block implementation (Mulgan & Albury, 2003; Edmond, et al., 2006; Meijers et al., 2006; Moussa, McMurray & Muenjohn, 2018; Smara, & Bogatyreva, 2023).
  - According to respondents, there is a lack of vision and objectives for RI in their organisations, which hinders implementation. One said: *There is no structured innovation in our organisation. (…) there are innovation procedures, but it is very incremental and slow. It is not based on a thought-out structured flow within the organisation, that is simply not present."* This confirms that vision plays an important role, which is concurrent with the literature (Young, Charns &

Shortell, 2001; De Vries, Bekkers, Tummers, 2014; Singh, Akbani & Dhir, 2020; Stelzl, Röglinger & Wyrtki, 2020)

- Environmental
  - Formal mandate: respondents mention that there is often not a formal mandate from the government to work on RI, or it is very limited. This means that all RI activities are done on top of the normal workload. Additionally, there are many different FROs who all have different mandates, and different speeds of implementation. This means that there is a lack of a harmonised approach. Even with a top down approach, this fragmentation could lead to a difficult implementation, as suggested by Boersma, Wagenaar and Wolbers (2012).

This article thus confirms the literature on the organisational level. However, on the individual and environmental levels, more research is needed, especially on the effect of ego and own initiative, as well as on the role of a formal mandate for RI. Literature highlights that in ambidextrous organisation, RI is part of the job description of individuals (Meijers et al., 2006; Owen, 2018; Brooks, Grugulis & Cook,2021; Koskinen, 2021). It however does not mention a formal mandate for FRO to work on RI.

Conceptual Model

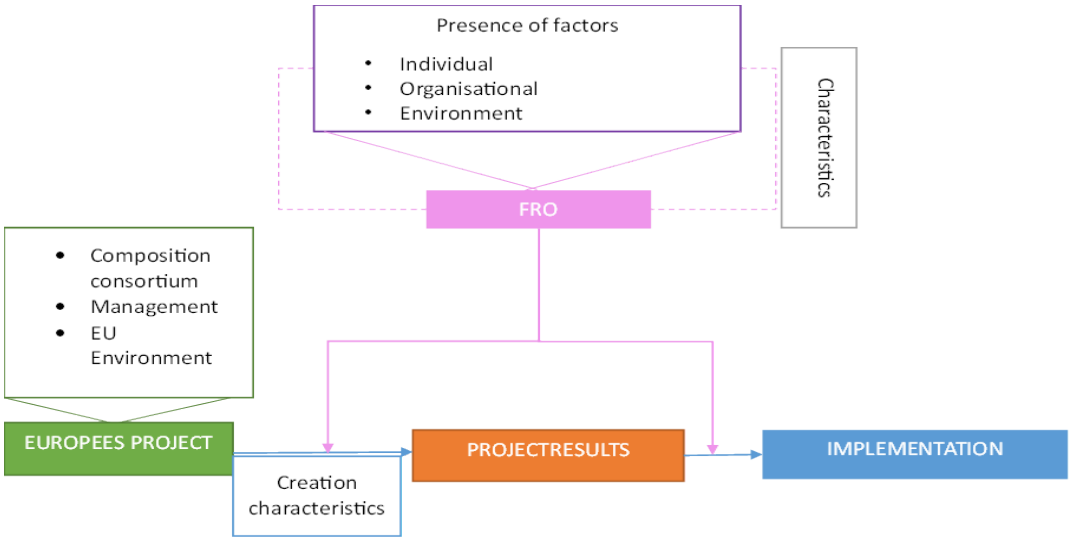


Figure 4. the renewed conceptual model

The above section has shown the role of the characteristics of EU project results, of the composition of the consortium the management, and the EU environment. This is visible in the new conceptual model in figure 4.

RQ1 asked: Why is there seemingly a failure to innovate of European Project results in FROs? It is clear from the respondents' answers that there can be a failure to implement project results. They often do not have desirable characteristics and thus, there is no enthusiasm to use them within FROs. The next step is determining why this is so. Is it the case because of the structure of EU projects? Or because of the structure of FROs? Or both? The respondents highlighted that the structure of the EU project can often be the problem . When the creation of characteristics goes wrong it leads to un-implementable project results. This happens during the EU project. It is however, also undeniable that there are factors in FROs that make implementation difficult to begin with. For example, a lack of vision and working without a structured flow of RI implementation.

Handbook

To give FROs a toolkit to prevent failure to innovate, and to answer RQ2, a handbook was created. This handbook was examined by all respondents and exists out of the following chapters (translated handbook: https://zenodo.org/records/10964582).

- Chapter 1 and 2: what are EU projects?
  - An explanation about the creation of the handbook, and of the most used jargon in EU projects.
- Chapter 3: requirements
  - The requirements that are needed for FROs to participate in EU projects. Of these requirements, a mix of people with project management experience and operational experience were

emphasized by the respondents. One respondent said: *"It is about bringing an operational background, the terrain knowledge to a more scientific or strategic project. You cannot get that knowledge from people who only know the procedures."*

- Chapter 4: selecting calls
  - Information on how to select calls. Respondents emphasised the importance of choosing project calls based on the concrete vision and objectives of the FRO. This chapter also gives an explanation on how to read project calls, based on Rösner (2023).
- Chapter 5: the proposal phase
  - During this phase, respondents mentioned that FROs should be aware of the composition of the consortium and of the coordinator. *"You need to look at: what is this project about, and how are we going to realise that with this group of people? That it is not all the same type of people when the call asks for someone else"*. Additionally, respondents mention that practical considerations need to be kept in mind.
- Chapter 6: the execution phase
  - Many respondents mention that during the execution phase, it is important to keep in mind your own vision and objectives and to be able to communicate those to the project partners.
- Chapter 7: the result phase
  - According to the respondents, this phase is crucial, but also the most overlooked. This is the time period that determines if results get implemented or not.
- Attachment: work pages
  - These can be filled in by FROs before, during and after the project, allowing them to track the project from beginning to end.

Recommendations

Based on the findings of this article, a few tentative recommendations can be made. Firstly, the current structure of EU projects includes hindrances that can lead to failure to innovate for FROs. These hindrances should be reduced. As such, the administrative burden of EU projects should be lowered if possible. On top of that, the current requirements demanded by the EC (high TRL for IA, required participation of end user organisation and specific dissemination plans) appear insufficient to ensure implementation in FROs. The utility of these requirements should be further evaluated. Secondly, FROs should be better prepared to participate in EU projects. They should develop a vision with concrete objectives about research and innovation implementation in their organisation. This should also explicitly address the project they are participating in. Thirdly, the gap between academics and FROs should be further closed by "translating" academic literature into concrete solutions, so that the project results are less abstract.

An overarching organisation that specializes in EU projects that employs active first responders and disseminates project results in different FROs could be a solution that addresses most of the recommendations. There would be expertise on the EU environment in this organisation, they would be prepared to participate, and they could support academics in the "translation" process.

Limits and Further Research

This was exploratory research only. More research is needed to further examine the findings of this article. A study with more FRO respondents in multiple European countries could be a good starting point. Additionally, how to encourage implementation of results in FROs should be further researched. For example, by exploring which additional requirements for EU projects could lead to more implementation.

There are several limitations to this research. Firstly, there is the possibility for a negativity bias, the tendency of respondents to focus more on the negative aspects then on positive aspects during an evaluation (Ito, Larsen, Smith & Cacioppo, 1998). Secondly, this article is focussed only on the FROs participating in European projects, it thus excludes the point of view of researchers, commercial parties, and the European Commission. Thirdly, the article had a small sample size, used the snowball method and was geographically in Belgium and the Netherlands. This sample therefore might not be wholly representative of all FROs participating in EU projects.

REFERENCES

Alexander, D. (2015) Disaster Risk Reduction and Resilience Strengthening the Links between Academics and Practitioners. *Third academic summit*

Allen, J. D., Towne, S. D., Maxwell, A. E., DiMartino, L., Leyva, B., Bowen, D. J., ... & Weiner, B. J. (2017). Measures of organisational characteristics associated with adoption and/or implementation of innovations: a systematic review. *BMC health services research*, 17, 1-10.

Akıncı, G., Alpkan, L., Yıldız, B., & Karacay, G. (2022). The link between ambidextrous leadership and innovative work behavior in a military organisation: The moderating role of climate for innovation. *Sustainability*, 14(22), 15315.

Birkinshaw, J. M., & Gibson, C. (2005). The ambidextrous organisation. Birmingham, UK: *Advanced Institute of Management Research*.

Boersma, K., Wagenaar, P., & Wolbers, J. (2012). Negotiating the 'Trading Zone'. Creating a Shared Information Infrastructure in the Dutch Public Safety Sector. *Journal of Homeland Security and Emergency Management*, 9(2), 1547-7355.

Brooks, J., Grugulis, I., & Cook, H. (2021). Unlearning and consent in the UK Fire and Rescue Service. Human Relations

Büschgens, T., Bausch, A., & Balkin, D. B. (2013). Organisational culture and innovation: A meta-analytic review. *Journal of product innovation management*, 30(4), 763-781.

Clausen, T. H., Demircioglu, M. A., & Alsos, G. A. (2020). Intensity of innovation in public sector organisations: The role of push and pull factors. *Public Administration*, 98(1), 159-176.

Cunningham, J. A., O'Reilly, P., Hooper, D., Nepelski, D., & Van Roy, V. (2020). The Role of Project Coordinators in European Commission Framework Programme Projects.

De Vries, H.A., Bekkers, V.J.J.M., Tummers, L.G. (2014). Innovations in the Public Sector: A Systematic Review and Future Research Agenda. Ottawa: *IRSPM conference*

Edmond, T., Megivern, D., Williams, C., Rochman, E., & Howard, M. (2006). Integrating evidence-based practice and social work field education. *Journal of Social Work Education*, 42(2), 377-396.

European Commission, Directorate-General for Research and innovation (2018). Horizon Europe impact assessment- *staff working document*. Part 1/3., Publications Office of the European Union. https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications/all-publications/horizon-europe-impact-assessment-staff-working-document_en

European Commission, Directorate-General for Budget, (2021a). The performance framework for the EU budget, under the 2021-2027 multiannual financial framework . Volume I, *Communication*, Publications Office of the European Union. https://data.europa.eu/doi/10.2761/573102

European Commission, (2021c), Horizon Europe Work programme 2021 -2022. general annexes. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/wp-call/2021-2022/wp-13-general-annexes_horizon-2021-2022_en.pdf

European Commission, European Innovation Council and SMEs Executive Agency, Scherer, J., Weber, S., Alveen, P. (2022a). European IP Helpdesk : successful valorisation of knowledge and research results in Horizon Europe : boosting the impact of your project through effective communication, dissemination and exploitation, Publications Office of the European Union. https://data.europa.eu/doi/10.2826/437645

European Commission (2022b), Horizon Europe programme, standard application form (HE RIA, IA). Version 6.0. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/af/af_he-ria-ia_en.pdf

European Commission (2022c), EU Funding & Tenders online manual. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/om_en.pdf

European Commssion, (2022d). HORIZON europe (HORIZON) Euratom Research and Training Programme (EUATOM) feneral model grant agreement, EIC accelerator contract. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/agr-contr/general-mga_horizon-euratom_en.pdf

Fu, L., Liu, Z., & Liao, S. (2018). Is distributed leadership a driving factor of innovation ambidexterity? An empirical study with mediating and moderating effects. *Leadership & Organisation Development Journal*, 39(3), 388-405.

González Fernández, S., Kubus, R., & Mascareñas Pérez-Iñigo, J. (2019). Innovation ecosystems in the EU: Policy evolution and horizon Europe proposal case study (the Actors' perspective). *Sustainability*, 11(17), 4735

Hero, L. M., Lindfors, E., & Taatila, V. (2017). Individual Innovation Competence: A Systematic Review and Future Research Agenda. *International Journal of Higher Education*, 6(5), 103-121.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of personality and social psychology*

Jacobs, S. R., Weiner, B. J., Reeve, B. B., Hofmann, D. A., Christian, M., & Weinberger, M. (2015). Determining the predictors of innovation implementation in healthcare: a quantitative analysis of implementation effectiveness. *BMC health services research*, 15, 1-13.

Koskinen, L. 2021. A leap to the Horizon – How did we manage? Talk-verkkolehti, Puheenvuoroja

Meijers, J. M., Janssen, M. A., Cummings, G. G., Wallin, L., Estabrooks, C. A., & YG Halfens, R. (2006). Assessing the relationships between contextual factors and research utilization in nursing: systematic literature review. *Journal of advanced nursing*, 55(5), 622-635.

Mosadeghrad, A. M. (2015). Developing and validating a total quality management model for healthcare organisations. *The TQM Journal*, 27(5), 544-564.

Moussa, M., McMurray, A., & Muenjohn, N. (2018). A conceptual framework of the factors influencing innovation in public sector organisations. *The Journal of Developing Areas*, 52(3), 231-240

Mulgan, G., & Albury, D. (2003). Innovation in the public sector. Strategy Unit, Cabinet Office, 1(1), 40

Narayanan, A.,& Altay, N.(2021). Ambidextrous humanitarian organisations. *Annals of Operations Research*

O Reilly, C. A.,& Tushman, M. L. (2004). The ambidextrous organisation. *Harvard business review*,82(4),74-83.

Owen, C. (2018). How emergency services organisations can: and do: utilise research. *Australian Journal of Emergency Management*, 33(2), 28-33.

Rip, A. (2016). The clothes of the emperor. An essay on RRI in and around Brussels. *Journal of Responsible Innovation*, 3(3), 290-304.

Rösner, A. (2023). Introduction on EU-proposal writing. *DEFERM summer school presentation.*

Serwiak, S. (2022) Practitioners in the Eu research and innovation projects. DG HOME F2 - innovation and security research. *VR and Police SHOTPROS conferentie*. https://vrandpolice.eu/final-conference-and-results-of-the-shotpros-project/

Singh, S., Akbani, I., & Dhir, S. (2020). Service innovation implementation: a systematic review and research agenda. *The Service Industries Journal*, 40(7-8), 491-517

Smara, R., & Bogatyreva, K. A. (2023). Moving to ambidextrous organisation: a systematic literature review. *Вестник СПбГУ. Менеджмент*, 22(2).

Soeters, J. L. 6 Ambidextrous Military. *The Viability of Human Security*, 109.

Somech, A., & Drach-Zahavy, A. (2013). Translating team creativity to innovation implementation: The role of team composition and climate for innovation. *Journal of management*, 39(3), 684-708.

Starkweather, A. R., & Kardong-Edgren, S. (2008). Diffusion of innovation: Embedding simulation into nursing curricula. *International Journal of Nursing Education Scholarship* , 5(1).

Stelzl, K., Röglinger, M., & Wyrtki, K. (2020). Building an ambidextrous organisation: a maturity model for organisational ambidexterity. *Business Research*, 13, 1203-1230

Strejcek, G., & Theil, M. (2003). Technology push, legislation pull? E-government in the European Union. *Decision Support Systems*, 34(3), 305-313.

Young, G. J., Charns, M. P., & Shortell, S. M. (2001). Top manager and network effects on the adoption of innovative management practices: A study of TQM in a public hospital system. *Strategic management journal*, 22(10), 935-951.

Knowledge on demand: the future of decision support systems?

Yohann Chasseray
IMT Mines Albi, Universite´ de Toulouse
yohann.chasseray@mines-albi.fr

Aurélie Congès
IMT Mines Albi, Universite´ de Toulouse
aurelie.conges@mines-albi.fr

Audrey Fertier
IMT Mines Albi, Universite´ de Toulouse
audrey.fertier@mines-albi.fr

Anne-Marie Barthe-Delanoë
IMT Mines Albi, Universite´ de Toulouse
anne-marie.barthe@mines-albi.fr

WiP Paper – Open Track
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

ABSTRACT

The availability and quality of the information system supporting crisis management depends on the existence of up-to-date, comprehensive and reliable knowledge bases, despite the multidisciplinary nature of the information to be exchanged, the actions, and the decisions to be taken. This article proposes a method to benefit from pre-trained language models to automatically update these knowledge bases at both design time and run time. We offer a model-based extraction process and a dedicated automated validation process. These two contributions are illustrated and discussed around a case study: an emergency decision support system used to prepare for and respond to an unexpected mass gathering at an abandoned French aerial base.

Keywords

Decision Support System, Knowledge Base, Generative AI, Crisis Management, Information Management

INTRODUCTION

Disaster management is usually structured around four phases: prevention, preparedness, response, and recovery (Alexander, 2002; Altay & Green, 2006). During a crisis response, situational awareness is the key to efficient and effective management of the situation (Coche et al., 2021; Endsley, 2001). Decision-making processes rely on a good understanding of the situation. In the context of the ever-growing amount of available data coming from the crisis sites, crisis responders can rely on decision-support systems that are able to process it through previously acquired knowledge. These systems are very useful in offering visualization and analysis tools to increase the understanding of the situation, but to function, they rely on a knowledge base. This knowledge base needs to be filled before the crisis and updated every time the situation evolves. This is crucial to ensure that the model of the situation given to the crisis responders is always relevant and represents what is currently going on. If extracting information about crises during the preparation phase has become rather easy, the question is how can they (i) fit into the structure of the knowledge base and (ii) be integrated with pre-existing knowledge to feed the situation model and eventually result in relevant and efficient decisions. How can the knowledge base of decision-support systems be filled with information and knowledge, both before and during the crisis? There are several ways to fill a knowledge base with information before the crisis, but the urgency and unexpectedness of crises make it impossible to have fully and sufficiently complete knowledge bases. This means that a way to update and complete the knowledge base during the crisis, to keep up with the ever-changing situation, has to be found.

The task of knowledge base population then consists of finding, for the situation under configuration, all instances of metamodels concepts as well as the relations that link them. As the crisis evolves, both instances and relations can be modified or updated. Some new instances may appear (a new risk is arising), and existing instances may see their relations with other instances appear or evolve (a resource is no longer available in a given area, the arisen risk impacts one or several actors). This constantly evolving nature of crisis associated with the high amount of information to be integrated to render a relevant model of the situation justifies the need to automate the updating of previously evoked concepts. This leads to the problematic addressed in this paper: How can we keep knowledgebases up to date without overwhelming managers in nominal mode, before the crisis, and in emergency mode, during the crisis?

Section 2 presents the context of the study and feedback on the use of conversational AI to update an existing knowledge base at design time and run time. Section 3 presents a literature review of the existing solutions to automatically fill knowledge bases. Section 4 presents our proposal to keep our knowledge base up to date without overloading emergency managers. Finally, Section 5 concludes with the impact and benefits of our feedback and proposal.

CONTEXT: THE DECISION-SUPPORT SYSTEM R-IOSUITE

This section presents our feedback on the use of pre-trained language models to facilitate the instantiation of the knowledge base of a decision support system: R-IOSUITE. R-IOSUITE is an open-source software dedicated to emergency management Salatge´ and Rebie`re-Pouyade, 2023.

R-IOSUITE Before Conversational-AI

The decision-support system (DSS) R-IOSuite presented in Fertier et al., 2020b aims to help decision-makers and crisis responders solve the crisis most efficiently. It has four main functionalities to achieve that goal: (i) manually

populate a knowledge base and model a crisis situation (ii) automatically interpret data flows, real or simulated, to update their model according to the current situation (Interpretation functionality); (iii) deduce, orchestrate and supervise a process combining human or automatic tasks to react to the current situation (Deduction functionality); (iv) adaptation to adverse events (Adaptation functionality). All these functionalities depend on a knowledge base, where R-IOSuite stores and infers information. The information concerns both general, core, knowledge common to several crisis instances, as specific, application, knowledge describing one particular site (location, sensitive building or infrastructure nearby, geographical information...) and its available stakeholders (firefighters, policemen, trucks or ambulances...). It is structured by a meta-model dedicated to represent collaborative situations (Benaben et al., 2020): the objectives, actors, resources, or risks, and their undergoing relations.

There are several ways to fill the R-IOSUITE knowledge base: manually with experts in design time (approximately 80% of the knowledge), during the preparation phase, manually in run time by the emergency managers during the response phase (1%), automatically in run time during the response phase (Coche et al., 2021; Fertier et al., 2020a) via machine learning (text from tweets, images form tweets) or complex event processing (sensor data and open data) (19%).

To illustrate the manual and automatic update of a knowledge base in R-IOSUITE, we propose a case study on an unauthorized mass gathering in a decommissioned airbase, developed in the context of the PICADORS research project. This project is funded by the European Union through the Re´gion Haut-de-France. The unauthorized aspect of this use case enhances the unexpected side of crises and the need to find ways to fill the knowledge base during the crisis, as the situation changes. Several actors are involved in the resolution of this crisis: the firefighters, the police, the French Red Cross, and the Prefecture of Cambrais where the aerial base is located. These actors have skills and competencies that must also be stored in the knowledge base so that the DSS can use that information to attribute the process' tasks.

R-IOSUITE Using Conversational AI to Fill Its Knowledge Base

Another possibility to fill the knowledge base would be to directly use conversational AI. This possibility can be used before the crisis, in the preparation phase when the knowledge base is filled with the available resources and concepts, but also during the crisis if the situation evolves and must be updated quickly. This second option is particularly interesting because it allows to enrich the knowledge base quickly, without having to look for information in operation plans or documents which can sometimes be hard to find. In the following subsection, we describe these two possible uses.

Filling the knowledge base before the crisis

Many of the actors involved in crisis management are the same from one crisis to another. This means that it is possible to anticipate the crisis by filling the knowledge base with information that is relevant in any kind of situation: for instance, a firefighter can extinguish a fire, and that is always true. To fill the knowledge base with this kind of information, it can be done manually, by automatic information extraction such as presented in the previous section, or with conversational AI.

In R-IOSuite, the knowledge base is filled with concepts such as roles that have functions: the role "firefighter" has functions such as "extinguish a fire" or "rescue people". It is possible to define these roles and functions either manually, or by requesting a conversational AI such as Chat GPT or Google Bard. In this case, the AI responds with several propositions that must be validated by the users of the DSS.

In the PICADORS use case, we asked the AI to help us find the functions of the actors from the Red Cross. It resulted in a list of possible functions such as treating injuries or sorting casualties. Among this list of functions, we were able to choose the ones that we thought were relevant to the use case and store them in the knowledge base.

Filling the knowledge base during the crisis

One of the characteristics of a crisis is its unexpected aspect. It is impossible to foresee every possible evolution of the situation and to be prepared for every possibility. Thus, it is necessary to be able to quickly modify or complete the knowledge base to keep it up-to-date. There again, it can be done manually, or with information extractions, but in some cases, it might be useful to use conversational AI.

In the PICADORS use case, we can imagine that the evolution of the situation requires at some point that some new actors get involved. If this evolution is unforeseen, the functions and skills of these new actors are probably missing from the knowledge base, which means that the DSS is unable to involve them in its response process proposal. Thus, we need to complete the knowledge base before asking R-IOSuite to deduce a new response process relevant to the new situation. Thanks to conversational AI, it is possible to quickly find the functions corresponding to the new actors and integrate them into the knowledge base if they are judged relevant by the decision-makers.

Benefits and Limits of Conversational AI to Fill an Existing Knowledge Base

The use of conversational AI-enabled us to reduce the time needed to implement the use case in R-IOSUITE. It also enables us to complete a knowledge base to cope with unexpected events, requiring specific knowledge that we wouldn't have thought of at the design stage. However, superfluous knowledge is added (noise) in the same way as missing knowledge, and AI cannot ensure the authenticity of the knowledge generated.

LITERATURE REVIEW ON DIFFERENT METHODS TO FILL THE KNOWLEDGE BASE

The aim of this section is to determine the extent to which a pre-trained linguistic model or machine learning models can be used to automate the updating of an information system's knowledge base, at both design and runtime.

Knowledge is available across an extensive collection of sources, ranging from structured data (databases, graphs, spreadsheets, etc.) to semi-structured (markup and JSON documents) and unstructured data (texts, such as PDF files, website pages, social media posts and comments, etc.).

Understanding, tagging, and structuring this large amount of data available from the web space into knowledge bases cannot be entirely achieved by hand, especially in the time-sensitive context of crisis management. Moreover, it is necessary to align the collected data with the existing knowledge structure on which the decision support systems rely (namely ontologies in the broad sense – as Gruber, 1993 originally defined it –, from OWL ontologies to knowledge graphs). The knowledge base population steps must then be automated to lessen the human effort and speed up the process of non-added-value tasks.

The literature review is based on the Scopus database, from 1980 to 2024, on the title, the abstract, and the author keywords fields. Table 1 presents the four parts of the query linked with the AND operator. It searches for publications about automated techniques of knowledge extraction, acquisition, population, feeding or collection of structured knowledge.

The first part of the query [1] looks for tools for structuring data in the broad sense (hence the use of keywords such as ontology, structure, base, etc.). The second part [2] seeks ways to get this knowledge automatically. The third part [3] focuses on the crisis management application domain. The fourth and last part of the Scopus query filters results by excluding domain areas that are not relevant to the research question (e.g., "VETE" for veterinary and "DENT" for dentistry).

The 56 papers returned by the query are filtered regarding their relevance with crisis domain and knowledge extraction concern. The 10 remaining related works are presented in Table 2 and evaluated through three criteria which are (1) the automated nature of the proposed extraction method, (2) the ability to refer to existing knowledge structure, (3) the complexity of extracted relations and (4) the type of processed data. The first column indicates the article reference. Columns 2 to 4 present our three criteria. Column 5 indicates which kind of data was used to perform knowledge extraction.

Table 1. The Scopus query

Operator	Number	Part of the query
	[1]	(base* OR database* OR structure* OR repositor* OR model* OR ontolog*) W/1 (knowledg* OR data OR fact* OR relation* OR entit* OR ontolog*)
AND	[2]	automat* W/1 (extract* OR acqui* OR populat* OR feed* OR collect*)
AND	[3]	cris* OR disast* OR "adverse event*" OR "unexpect* event*"
AND	[4]	EXCLUDE (SUBJAREA , "DENT") OR EXCLUDE (SUBJAREA , "VETE") OR EXCLUDE (SUBJAREA , "NURS") OR EXCLUDE (SUBJAREA , "CENG") OR EXCLUDE (SUBJAREA , "IMMU") OR EXCLUDE (SUBJAREA , "PHAR") OR EXCLUDE (SUBJAREA , "CHEM") OR EXCLUDE (SUBJAREA , "MATE") OR EXCLUDE (SUBJAREA , "ARTS") OR EXCLUDE (SUBJAREA , "PHYS") OR EXCLUDE (SUBJAREA , "EART") OR EXCLUDE (SUBJAREA , "BIOC") OR EXCLUDE (SUBJAREA , "MEDI")

Table 2. The results of the literature review organised according to our four criteria. ***Abbreviations : N.R Not Relevant; N/A Not available***

Reference	Automated method	Compliance	Complexity	Data
Angeles and Kijewski-Correa, 2022	No	Yes	Model specific	images, texts
Bayramog˘lu and Uzar, 2023	No	Yes	Taxonomic	images
Chasseray et al., 2021	Yes	Yes	Taxonomic	texts
Chaulagain et al., 2019	Yes	No	Taxonomic	texts
D'Amico et al., 2011	No	Specific	Non taxonomic	discussions
Guo et al., 2023	No	No	*N.R.*	images
Nguyen et al., 2011, 2012, 2013	Semi	Yes	Non taxonomic	tweets (semi-structured)
Venkatachalam et al., 2020	*N/A*	Yes	*N/A*	texts
Wang et al., 2020	Yes	Specific	Non taxonomic	texts
Zong et al., 2022	Semi	BDD	Non taxonomic	tweets (semi-structured)

A few related works extract territorial information that are relevant in case of a crisis based on aerial data. Bayramoğlu and Uzar, 2023 focus on the inventory of roads from unmanned aerial vehicle images. Similarly, Guo et al., 2023 identify cropping zones that have been affected by bad weather conditions and are susceptible to crop lodging. They use satellite image and image processing techniques to identify the concerned zones. In the same area of research, Park et al., 2015 build topographic landscape representation, and notably tree height for the management of disasters such as forest fires. As techniques presented in these articles are relevant in term of information extraction, they do not involve the population of an existing knowledge structure or are not related to a decision support system. Bayramoğlu and Uzar, 2023 is the only one proposing an ontology built on 4 classes to reference extracted elements such as Bicycle (road), Sidewalk, Road and Marking. However, these classes are not pre-existing the extraction method and are designed on purpose based on the extraction task results.

Angeles and Kijewski-Correa, 2022 couple image sourced information extraction with semi-structured textual data such as tax assessor data or building permits. Their framework results in a model characterizing the sensibility of buildings. This model can be reused in case of a crisis or disaster inducing high winds. However, as no pre-existing knowledge structure is used, the created models can't be easily integrated in a more global crisis management system.

More generally, Babic˘ et al., 2022 list a wide group of ontology learning tools that can be used in the case of crisis management for the building of a knowledge structure. These tools are nevertheless not relevant when dealing with an existing knowledge structure such as a metamodel or an ontology which has to be populated and not rebuilt from scratch.

Nguyen et al., 2011, 2012, 2013 propose the definition of earthquake semantics networks, that can be populated from tweet analysis and further integrated in a crisis management system. As such a knowledge management system ensures interoperability with crisis management systems, the population of the expected knowledge graph from tweets requires prior training steps and associated training datasets which makes the approach not fully automated.

Some automated extraction methods for crisis have been proposed in the literature. Chaulagain et al., 2019 use named entity recognition and the definition of specific patterns for the extraction of crisis-related components

(number of deaths, injuries, date and locations). Similarly, Chasseray et al., 2021 apply their domain-independent ontology population patterns to a crisis metamodel, for the extraction. As the proposed methods in these works are both automated, they focus only on taxonomic relations (extraction of concept-instance relations) and do not tackle the domain-specific relations that may occur between two instances of a knowledge base. **Is there a possibility to benefit from pre-trained linguistic models, while limiting the generation of noise and optimizing the authenticity of the generated knowledge?**

PROPOSAL

As natural language processing can be used to populate a knowledge base, their performances are highly dependent on the domain and considered case study. This section discusses the possibility of designing extraction systems encompassing a larger amount of information. As enlarging the volume of extracted information also supposes a rise in irrelevant information, a proposition is made to pair the extraction systems with a conversational AI that could be used as an automated validation tool in order to filter the generated noise.

Natural Language Processing for Relation Extraction

Considering the fact that exploited data is of a textual nature, the field of Natural Language Processing provides a large pallet of tools that can be mobilized to extract knowledge based on unstructured data. Among all tasks of natural language processing, two are relevant for knowledge base population which are instance extraction and relation extraction.

Instance extraction identifies in textual data, pieces of text that can semantically be linked to an existing concept as an instanced version of this concept. This is especially used in ontology and knowledge graph population, as it allows, from a set of defined concepts, to extend some of them with real-world instances.

Relation extraction is a process that aims at extracting semantic meaning linking to parts of a given text. Commonly, the parts to be linked through a relation correspond to instances of previously defined concepts (author-artwork, country-capital city, athlete-discipline).

Our illustrative case study focuses on the extraction of relations between role instances (Firefighter, Caregiver, Policeman, Military person, RedCross, Institution, Officials, Security agent) and function instances (Establish a perimeter on site, Extinguish a fire, Launch an alert, Heal a victim, Drive an engine). We state that instead of extracting instances and relations in separate processes, it is either possible to use the former to extract the latter or to extract both in the same process.

In our specific case, a scenario can be imagined where the objective is to know which Role is involved in the crisis and which Function this role has (what is a policeman able to do in case of a fire ?). This need exists both during the preparation phase and response phase, as associations between a Role and a Function may also appear during a crisis, due to its evolution, if a Role sees itself granted with new functions for example.

Knowing that RedCross is a role, and Protect people is a Function, the goal of the first approach is to identify all occurrences of the RedCross instance and Protect instance in a given text and see if they can be related in the text. This approach is relatively direct, but supposes that instances are already known.

When, none of these instances is identified however, another approach could be to drive the extraction with information contained by the semantic meaning of the relation linking the two concepts. Knowing the type of relation that we want to extract is a great advantage regarding the diversity of textual data. The proposed rule-based approach, presented in Figure 1, hence relies on the richness of verbal relation in textual data and the information they provide to identify instances. For each relation identified in the metamodel, the extraction process can be decomposed into three main steps:

1. Identifying all active verbal forms that may traduce the relation identified between the two concepts of interest. In the case of Role-Function relations, this list can contain verbal forms such as *Is able to*, *Can*, *has to*, *acts to*, *Is tasked with*, *Is capable of*, for example all traducing the ability of a Role to operate a task (Function).
2. Based on the identified forms, build all possible triples involving one of the verbal forms identified in the text and extract the pieces of text that the verbal form links using a predefined extraction pattern structure.
3. Validate the triples and register them in the knowledge base linking each piece of text to its attributed concept (attributed concepts are known as they are linked to the relation in the metamodel).

This approach is relevant as it does not need any pre-training on data specifically related to the crisis domain, as would a language model dedicated to crisis management area for example.

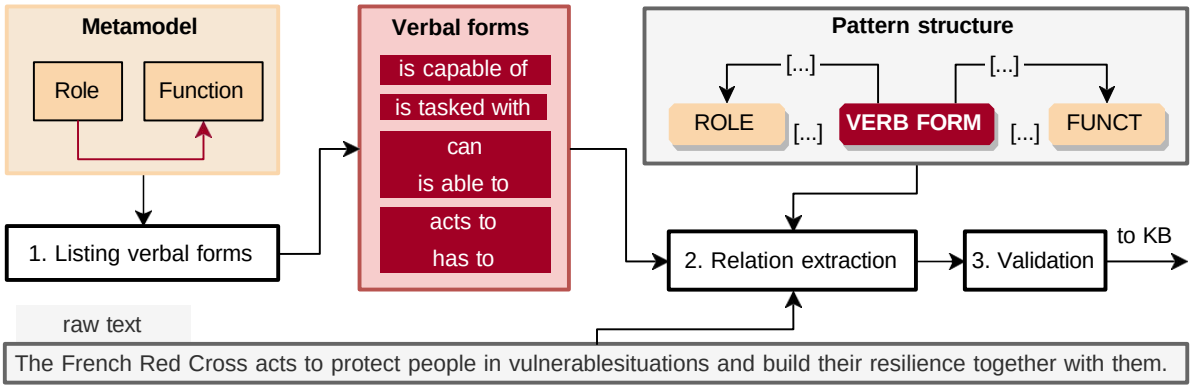


Figure 1. Proposed pattern-based extraction process of Role-Function relations

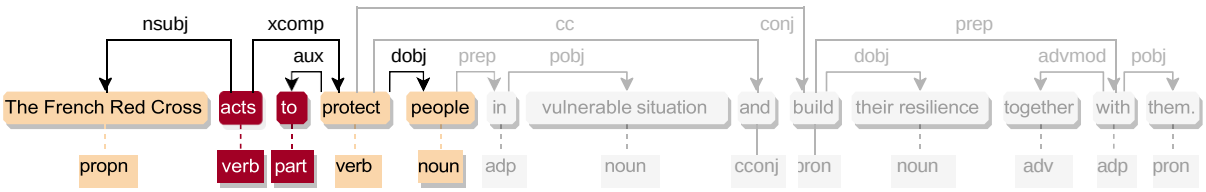


Figure 2. Application of a restrictive Role-Function extraction pattern

Defining Covering and Generic Extraction Pattern

Step 2 of the proposed approach can be performed through the application of extraction patterns. A main limitation of extraction patterns however is that even though they are very precise when specifically defined, they can miss a lot of information. If this information is expressed through a slightly different expression, a pattern that previously worked and extracted the wanted information is no longer effective. Loosen the pattern, at the opposite, tends to greatly extend the range of covered possibilities, but in counterpart, may lead to the extraction of a lot of irrelevant or false information.

Figure 2 and Figure 3 show two examples of extraction patterns being respectively restricted to direct expression of a function (figure 2) and extended to encompass more possibly valid relations 3. In the represented sentence : *The French Red Cross acts to protect people in vulnerable situations and build their resilience together with them.* Two Role-Function relations can be identified and should be extracted which are the relation between the Role *Red Cross* and the Function *Protect people* and between the Role *Red Cross* and the Function *build their resilience*.

To extract these relations, a possibility is the definition of extraction patterns, adapting the definition of patterns proposed by Chasseray et al., 2023, which allows the generic extraction of concept-instance relations so that they can extract relations based on previously defined verbal forms.

First pattern (2) looks for the restricted extraction of the following sequence of tokens:

$$\text{PROPN} \overset{\text{nsubj}}{\dashrightarrow} [\text{VERBAL FORM}] \overset{\text{xcomp}}{\dashrightarrow} \text{VERB} \overset{\text{dobj/pobj}}{\dashrightarrow} \text{NOUN}$$

Where :

1. the NOUN may later be defined as a Role,
2. the RELATION VERBAL FORM matches with one of the verbal form defined in step 1,
3. the VERB $\overset{\text{dobj/pobj}}{\dashrightarrow}$ NOUN sequence is used to identify a potential Function.

Using such a pattern on the example results in the extraction of one correct function in that context (*protect people*). This pattern can nevertheless be criticised as it misses the second function expressed in the sentence (*build their resilience*) as it is restricted to a limited group of expressed relations.

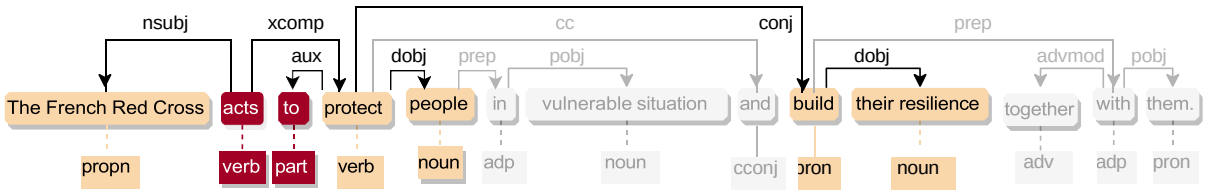


Figure 3. Application of a loose Role-Function extraction pattern

A larger extraction pattern looking for a more generic sequence of tokens:

$$\text{NOUN} \overset{\text{any}}{\dashrightarrow} \ldots \overset{\text{any}}{\dashrightarrow} [\text{VERBAL FORM}] \overset{\text{any}}{\dashrightarrow} \ldots \overset{\text{any}}{\dashrightarrow} \text{VERB} \overset{\text{dobj/pobj}}{\dashrightarrow} \text{NOUN}$$

allows the extraction of every candidate on each side of the matched VERBAL FORM, without paying attention either to the nature of syntactic dependencies linking the different elements to extract or the length of the dependency path linking them. In the provided example (see Figure 3), this pattern performs well as it is able to extract both targeted relations.

Contrarily to specific dependency patterns described in Chasseray et al., 2023, this extraction pattern is agnostic regarding the nature of the dependency links, which ensures a wider coverage and a larger amount of extracted information.

However, as relaxation of the extraction pattern can be seen as a solution to cover all needed information, the risk is flooding the relevant information in too much noise. When applied to larger textual content, a larger pattern might then include numerous relationships, some of which may not be semantically correct, leading to an accumulation of incorrect triples, that will surely extend the validation task. In that case, manual validation of extracted triples is not reasonable, especially during the response phase, where deduced scenarios and taken actions should be proposed as fast as possible to decision-makers.

Assisting Validation Filtering with Conversational AIs

In the methodology proposed, the third step is crucial to ensure that the knowledge base is provided with information that is correct, and sufficient for scenario deduction. This inherent nature of pattern extraction is in conflict with the objective of a complete model of the situation, by stating that too precise patterns will not extract all the available information, resulting in an incomplete model.

Despite the fact that conversational AI outputs are raising questions concerning their relevance or the lack of sources from which they suffer, they can be used as a tool to validate already extracted (and sourced) knowledge.

Directly asking a large language model and its associated conversational AI to populate a knowledge base remains questionable depending on scenarios and use cases for two main reasons which are (i) the impossibility of tracking the origin of the extracted information and (ii) the risk to face inferred knowledge and add apparently realistic information, that does not exist in real life, or not in the current situation.

However, using such conversational AI as a post-extraction validator does not affect the whole knowledge base population system as it is provided with traceable information, and only states its validity. We propose a first algorithm flowchart diagram in 4 to automate validation using conversational AIs. The proposed methodology lies on predetermined verbal forms and extracted relations to generate different prompts for validation. Generated prompts are submitted to a conversational AI and validation outputs (True/False statement + confidence) are aggregated to establish a general True/False statement on the relation. Extracted relations that are stated as False by the conversational AI are ignored. Validated relations are however added to the knowledge base. Human validation is however still required to confirm the extracted relation as true and available for further use by a decision support system. This is mandatory, as conversational AIs and Large Language Models can falsely qualify an extracted relation as True. Implementation of such a knowledge extraction framework can be parameterised with a threshold stating when a relation should be considered True or False. This threshold could be set on the average rate of decision, but can also use the confidence indicator.

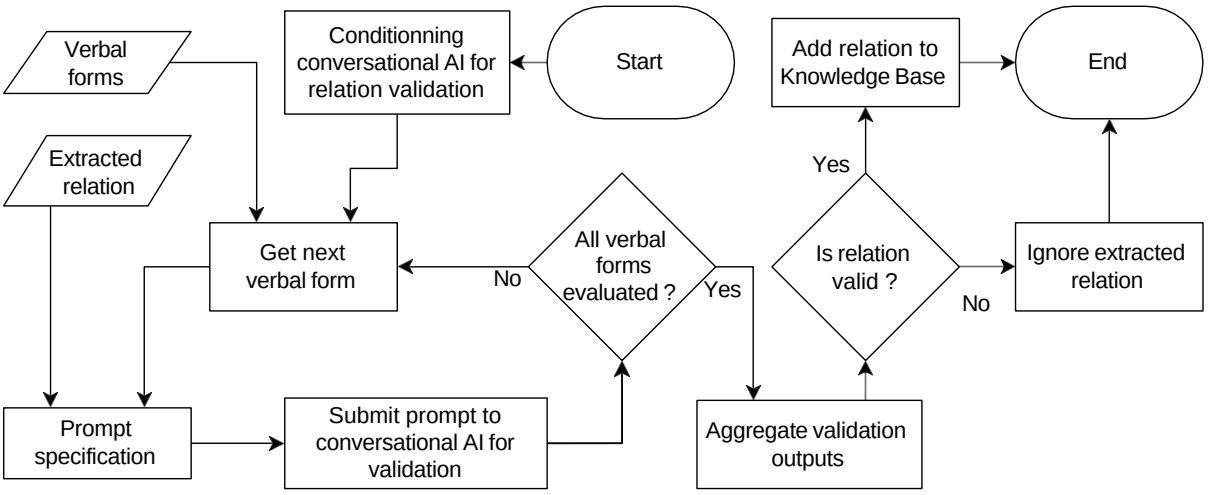


Figure 4. Flowchart diagram for the automated validation process of extracted relations

DISCUSSION ON IMPACT AND BENEFITS

The use of the methodology proposed could improve the ways to fill the R-IOSUITE knowledge base: automate part of the manual updates with experts in design time (at least 80% of the 80%), during the preparation phase; automate part of the unexpected manual updates during the response phase (at least 50% of the 1%); and improve the interpretation of texts or images from tweets, during the response phase; while limiting noise and optimizing the authenticity of the knowledge generated.

For the French fire department in the Pas-de-Calais department, our method of preparing and assisting a response to an unexpected mass gathering has enormous potential. This innovative approach could revolutionize crisis management by enabling pre-filling and automatic, controlled updating of the knowledge base of the emergency decision support system. What's more, in the event of a crisis, this system could prove invaluable in enabling the reconfiguration of prepared emergency plans. This would guarantee a more informed and effective response.

CONCLUSION

This paper presents four contributions: a feedback on the use of conversational AI to instantiate a knowledge base at design time and run-time; a literature review on existing systems able to automatically instantiate a knowledge base in the crisis management field; a method to extract instances and relations (patterns); and a method to automatically validate the extracted patterns.

The example provided in this article is detailed for illustration purposes. Future work on this subject could consist of systematic implementation of the proposed AI-assisted knowledge base population methods to apply it to several crisis cases and assess its adaptability to different use cases.

REFERENCES

Alexander, D. E. (2002). *Principles of emergency planning and management*. Oxford University Press on Demand.

Altay, N., & Green, W. G. (2006). OR/MS research in disaster operations management. *European Journal of Operational Research*, *175*(1), 475–493. https://doi.org/10.1016/j.ejor.2005.05.016

Angeles, K., & Kijewski-Correa, T. (2022). Data-driven framework for automated simulation of wind and windborne debris effects for hurricane regional loss estimation. *Journal of Wind Engineering and Industrial Aerodynamics*, *230*. https://doi.org/10.1016/j.jweia.2022.105167

Babicˇ, F., Buresˇ, V., Cˇ ech, P., Husa´kova´, M., Mikulecký, P., Mls, K., Nacha´zel, T., Ponce, D., Štekerová, K., Triantafyllou, I., Tucˇn´ık, P., & Zanker, M. (2022). Review of tools for semantics extraction: Application in tsunami research domain. *Information (Switzerland)*, *13*(1). https://doi.org/10.3390/info13010004

Bayramoğlu, Z., & Uzar, M. (2023). Performance analysis of rule-based classification and deep learning method for automatic road extraction. *International Journal of Engineering and Geosciences*, *8*(1), 83–97. https://doi.org/10.26833/ijeg.1062250

Benaben, F., Fertier, A., Montarnal, A., Mu, W., Jiang, Z., Truptil, S., Barthe-Delanoe¨, A.-M., Lauras, M., Mace-Ramete, G., Wang, T., Bidoux, L., & Lamothe, J. (2020). An AI framework and a metamodel for collaborative situations: Application to crisis management contexts. *Journal of Contingencies and Crisis Management*, *28*(3), 291–306. https://doi.org/10.1111/1468-5973.12310

Chasseray, Y., Negny, S., Barthe-Delanoe, A., & Le Lann, J. (2021). Automated unsupervised ontology population system applied to crisis management domain. In Adrot A., Grace R., Moore K., & Zobel C.W. (Eds.), *Proc. Int. ISCRAM Conf.* (pp. 968–981, Vol. 2021-May). Information Systems for Crisis Response; Management, ISCRAM.

Chasseray, Y., Barthe-Delanoe¨, A.-M., Ne´gny, S., & Le Lann, J.-M. (2023). Knowledge extraction from textual data and performance evaluation in an unsupervised context. *Information Sciences*, *629*, 324–343.

Chaulagain, B., Bhatt, B., Panday, S., Shakya, A., Newar, D., & Pandey, R. (2019). Casualty information extraction and analysis from news. In Franco Z., Gonzalez J.J., & Canos J.H. (Eds.), *Proc. Int. ISCRAM Conf.* (pp. 1002–1011, Vol. 2019-May). Information Systems for Crisis Response; Management, ISCRAM.

Coche, J., Kropczynski, J., Montarnal, A., Tapia, A., & Benaben, F. (2021). Actionability in a Situation Awareness world: Implications for social media processing system design. *ISCRAM 2021 - 18th International Conference on Information Systems for Crisis Response and Management*, (2391), p.994.

D'Amico, A., Buchanan, L., Goodall, J., & Walczak, P. (2011). Mission impact of cyber events: Scenarios and ontology to express the relationships between cyber assets, missions and users. *Eur. Conf. Inform. Manage. Eval.*, 388–397.

Endsley, M. R. (2001). Designing for situation awareness in complex systems. *Proceedings of the Second International Workshop on symbiosis of humans, artifacts and environment*, 1–14.

Fertier, A., Barthe-Delanoe¨, A.-M., Montarnal, A., Truptil, S., & Be´naben, F. (2020a). A new emergency decision support system: The automatic interpretation and contextualisation of events to model a crisis situation in real-time. *Decision Support Systems*, *133*, 113260.

Fertier, A., Barthe-Delanoe¨, A.-M., Montarnal, A., Truptil, S., & Be´naben, F. (2020b). A new emergency decision support system: The automatic interpretation and contextualisation of events to model a crisis situation in real-time. *Decision Support Systems*, *133*, 113260. https://doi.org/10.1016/j.dss.2020.113260

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, *5*(2), 199–220.

Guo, R., Zhu, X., & Liu, T. (2023). Automatic detection of crop lodging from multitemporal satellite data based on the isolation forest algorithm. *Computers and Electronics in Agriculture*, *215*. https://doi.org/10.1016/j.compag.2023.108415

Nguyen, T.-M., Kawamura, T., Tahara, Y., & Ohsuga, A. (2012). Building a timeline network for evacuation in earthquake disaster. *AAAI Workshop Tech. Rep.*, *WS-12-13*, 15–20.

Nguyen, T.-M., Kawamura, T., Tahara, Y., & Ohsuga, A. (2013). Toward information sharing of natural disaster: Proposal of timeline action network. *Commun. Comput. Info. Sci.*, *358*, 145–157. https://doi.org/10.1007/978-3-642-36907-0 10

Nguyen, T.-M., Koshikawa, K., Kawamura, T., Tahara, Y., & Ohsuga, A. (2011). Building earthquake semantic network by mining human activity from Twitter. *Proc. - IEEE Int. Conf. Granular Comput., GrC*, 496–501. https://doi.org/10.1109/GRC.2011.6122647

Park, W.-Y., Sohn, H.-G., & Heo, J. (2015). Estimation of forest canopy height using orthoimage-refined digital elevation models. *Landscape and Ecological Engineering*, *11*(1), 73–86. https://doi.org/10.1007/s11355-013-0238-3

Salatge´, N., & Rebie`re-Pouyade, S. (2023). R-IOSuite V2023.08.01 [Computer software].

Venkatachalam, S., Subbiah, L., Rajendiran, R., & Venkatachalam, N. (2020). An ontology-based information extraction and summarization of multiple news articles. *International Journal of Information Technology (Singapore)*, *12*(2), 547–557. https://doi.org/10.1007/s41870-019-00367-x

Wang, X., Gui, D., Li, H., & Gui, H. (2020). Automatic Construction of Coal Mine Accident Ontology. In Abawajy J.H., Choo K.-K.R., Islam R., Xu Z., & Atiquzzaman M. (Eds.), *Adv. Intell. Sys. Comput.* (pp. 1366–1374, Vol. 1017). Springer Verlag. https://doi.org/10.1007/978-3-030-25128-4 168

Zong, S., Baheti, A., Xu, W., & Ritter, A. (2022). Extracting a Knowledge Base of COVID-19 Events from Social Media. In Calzolari N., Huang C.-R., Kim H., Pustejovsky J., Wanner L., Choi K.-S., Ryu P.-M., Chen H.-H., Donatelli L., Ji H., Kurohashi S., Paggio P., Paggio P., Xue N., Kim S., Hahm Y., He Z., Lee T.K., Santus E., . . . Na S.-H. (Eds.), *Proc. Main Conf. Int. Conf. Comput. Linguist., COLING* (pp. 3810–3823, Vol. 29). Association for Computational Linguistics (ACL).

Future AI in crisis management: Proposing a bio-inspired, neuro-symbolic architecture

Aurélie Montarnal
IMT Mines Albi
aurelie.montarnal@mines-albi.fr

Frédérick Bénaben
IMT Mines Albi
frederick.benaben@mines-albi.fr

Rob Grace
Texas Tech University
rob.grace@ttu.edu

Grégoire Borst
LaPsyDÉ - UMR CNRS 8240, Universite´
Paris Cite´
gregoire.borst@parisdescartes.fr

Paul Gaborit
IMT Mines Albi
paul.gaborit@mines-albi.fr

ABSTRACT

This paper addresses the evolving relationship between Artificial Intelligence (AI) and crisis management, spanning over a decade. Initially sparked by social sciences' interest and facilitated by abundant data, particularly from social media, the collaboration between these fields has streamlined decision-making processes. However, the distinctive, uncertain nature of crises presents challenges in adapting AI to varying contexts. In exploring existing frameworks like Common Operational Picture (COP), Situational Awareness (SA), and Sensemaking, the paper finds a potential misalignment between recent AI architectures, predominantly symbolic or neural, and these established frameworks. To bridge this gap, the paper proposes a bio-inspired neuro-symbolic AI architecture, emphasizing its application at the Sensemaking level during crisis data exploitation. Leveraging insights from neurosciences advancements, the paper aims to enhance the adaptability and effectiveness of AI systems in crisis situations.

Keywords

Situation Awareness, Sensemaking, neuro-symbolic AI, neurosciences

INTRODUCTION

For more than a decade, the crisis management community has been interested in using Artificial Intelligence (AI) to support decision-makers and their information needs (Palen et al., 2020). This enthusiasm first came from the social sciences, whose studies were historically prevalent in the study of past crisis management, and naturally joined the field of AI thanks to easier access to voluminous streams of data - notably social media - that can be exploited computationally(Palen & Anderson, 2016). The literature shows that the two fields have been able to work together quickly and effectively, and the richness of the inter-disciplinary collaborations set in motion on this occasion no longer needs to be demonstrated, particularly within the ISCRAM community (Reuter et al., 2018).

With this in mind, the last few years have seen a strong interest in the application of AI to quite specific and not always generalizable tasks. While the basic issue remains that of saving crisis managers time and resources in accessing and exploiting rich, heterogeneous data flows, the challenges are not limited to this. The context of a crisis is unique in that it is by definition uncertain, sometimes unknown, and each time different in the course of events.

Slam et al. (2015) summarizes three challenges for designing a decision-support system for crisis management: represent knowledge, reason about knowledge in real-time, and adapt knowledge to unknown situations. Knowledge

representation capability refers directly to the decision support system's ability to process data flows in real time in order to establish an up-to-date map of the crisis context and its events. The systems affected by crises are generally complex, and the decision-making process, in its first stage of understanding the situation, is just as complex, involving a wide range of expertise and points of view. AI, with its ability to deal with large data flows and apply logical rules, is a strong ally to promptly address this first information need problem in decision support systems. In the literature, this ability of a decision support system to process data (human or machine) in order to understand a crisis situation is generally grouped under three terms, each with specific levels of processing and complexity: *Common Operational Picture (COP)*, *Situational Awareness (SA)* and *Sensemaking*. Although the definitions were clearly formulated a few years ago, the work proposed in AI, influenced since then by a remarkable progress in technology, has sometimes been developed independently.

While most AI architectures developed for crisis management are based almost exclusively on symbolic or neural architecture, neuro-symbolic architectures seem more likely to solve certain challenges, especially at the Sensemaking level. Yet, neuro-symbolic AI architectures, beside several attempt, still do not meet the level of expectation of a Sensemaking-level AI support. In parallel, neurosciences have undergo a strong progress in the last years and human brain cognition processes are now better known, providing promising inspiration for future AIs.

In this context, this paper serves two purposes. The second Section of this paper questions the current alignment between state-of-the-art AI architecture and the information needs of crisis responders formulated in COP, SA, and Sensemaking literatures. Hence, it focuses on a litterature review of the three formulations of informational needs and discusses the underlying processes for information extraction by individual or (inter-)organizational crisis responders. The third Section proposes a bio-inspired neuro-symbolic AI architecture to address the Sensemaking-level challenges in the exploitaiton of data to extract crisis comprehensive situation models. Finally, the conclusion will provide an opportunity to take a closer look at the current limits of the approach and future work.

LITERATURE REVIEW

Informational Needs and Human Processing

Common Operational Picture

The aim of a COP, which is generally the heart of Command and Control Systems, is to provide decision-makers with a *single* perceived and shared situational picture (Kuusisto et al., 2005). In essence, the SA-based decision-making paradigm involves the actor utilizing their knowledge to make decisions based on their perception, understanding, and projection of the dynamic environment they are operating in. On the other hand, COP serves as an information-sharing tool, allowing for the unambiguous modeling of a situation and therefore facilitating a fluid communication between actors.

Situational Awareness

SA is generally associated with Endsley (1995)'s famous framework, which involves perceiving, understanding, and projecting a situation as a reliable basis for decision-making. First, Endsley unambiguously positions SA as a state of knowledge, and refers to the process of accessing this state as *situation assessment*, enabling to "achieve, acquire and maintain SA". Second, in SA, Endlsey includes knowledge of the state of a dynamic environment and explicitly excludes any static knowledge or expertise. While the boundaries of this dynamic state of knowledge of a system in motion are clear, they gain in complexity in "real life", where decisions reflect coordination and even collaboration between several actors, just as in crisis management. Here again, Endsley's framework proposes a definition of *team SA*, i.e. a complete SA of each individual concerning his or her responsibilities, with areas of overlap between actors being acquired individually or through information sharing.

According to this definition of a state of knowledge necessary for decision-making, the acquisition process itself underpins cognitive mechanisms specific to human reasoning. This definition aligns with the broader understanding of reasoning by Leighton and Sternberg (2004) : "reasoning is broadly defined as the process of drawing conclusions. [...] these conclusions inform problem-solving and decision-making endeavors", thus also unifying goal-oriented paradigms, which in turn require the implementation of an attention mechanism on the part of the actor. Although Flach (2015)'s argument is probably beyond the scope of this paper, he stresses the importance of thinking about SA within a contextualized framework in which "meaning is defined in terms of functional significance or utility", and questions the status or rather the influence of sociotechnical systems ("thinking machines") in this framework.

SA can be realized through the individual access to *actionable information*. Actionable information is defined by Zade et al. (2018) as " the right information [that] reach the right person at the right time". Coche et al. (2021) operationalize the actionable information under four criteria: relevant, timely, precise and reliable. Therefore, the automation of actionable information mining comes as a crucial challenge to support SA with intelligent systems.

Sensemaking

Finally, the process of Sensemaking, although variously defined in the literature (Klein et al., 2007; Pirolli & Card, 2005), refers to processes "by which people seek plausibly to understand ambiguous, equivocal or confusing issues or events" (Brown et al., 2015). The term "Sensemaking" originates from Organizational Studies, but there is no consensus in the literature regarding its mechanism (Brown et al., 2015; Muhren & Van de Walle, 2010; Weick et al., 2005). The notion of the collective is called into question, as Sensemaking may be understood as an individual and/or collective process. In the context of crisis management, Weick's work in favor of a collective Sensemaking process is best summed up in the 7 properties summed up in the following citation: "Taken together these properties suggest that increased skill at Sensemaking should occur when people are *socialized* to make do, be resilient, treat constraints as *self-imposed*, strive for *plausibility*, *keep showing up*, use *retrospect* to get a sense of direction, and *articulate descriptions* that energize. These are micro-level actions. They are *small actions*. But they are small actions with large consequences." (Weick et al., 2005).

On COP, SA and Sensemaking Integration

As a conclusion of this section, we demonstrate that comprehending a situation, a crucial step towards decision-making in a dynamic environment, must be considered at three interconnected levels: (i) the cognitive processes that underpin human reasoning; (ii) the social context that influences individual evolutions as much as individuals influence social behavior; and (iii) the importance of a common language with a moderate level of interpretation, so as to favor non-ambiguous shared modeling.

Moreover, an AI to support Sensemaking, in its collective process definition, should address a two-dimensional challenge:

- *horizontal dimension* - the expertise and type of data to be processed fluctuate according to the type of crisis and the specific context;
- *vertical dimension* - the granularity of the information extracted from the data varies directly with the level of crisis (e.g. regional or national) and the decision-making level of the decision-maker who has the AI tools in hand (Benaben et al., 2016).

Artificial Intelligence to Meet Informational Needs

The field of crisis management, like many other fields in recent years, is showing strong growth in the use of AI, as can be seen from Figure 1, representing the number of publications on Scopus in response to the search "('crisis management' or 'emergency management') AND 'artificial intelligence'".

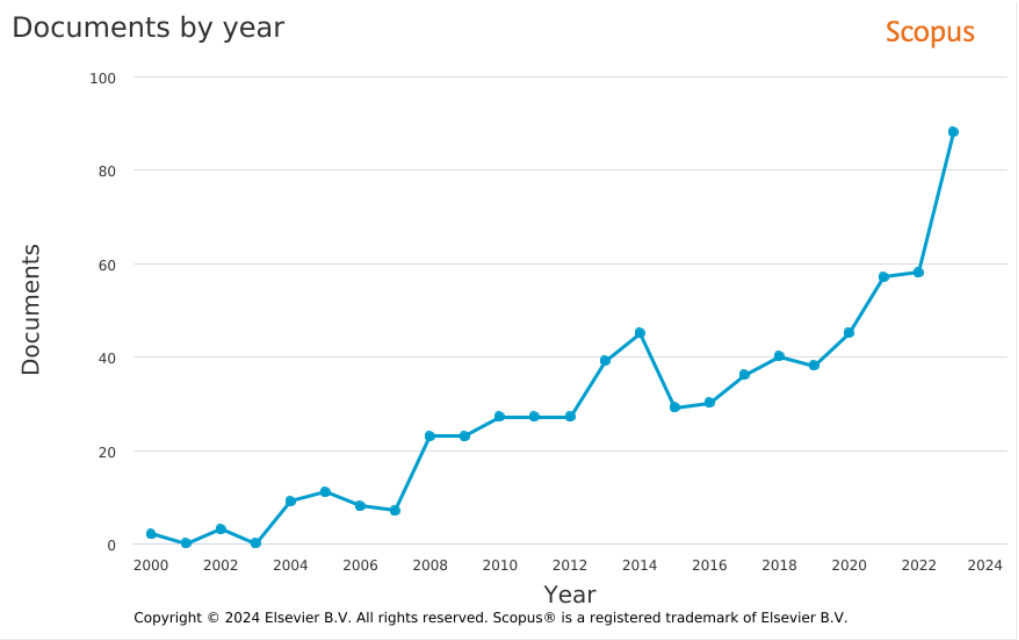


Figure 1. Evolution of the number of published paper on the Scopus scientific database, search string *TITLE-ABS-KEY ("artificial intelligence" AND ("crisis management" OR "emergency management"))*.

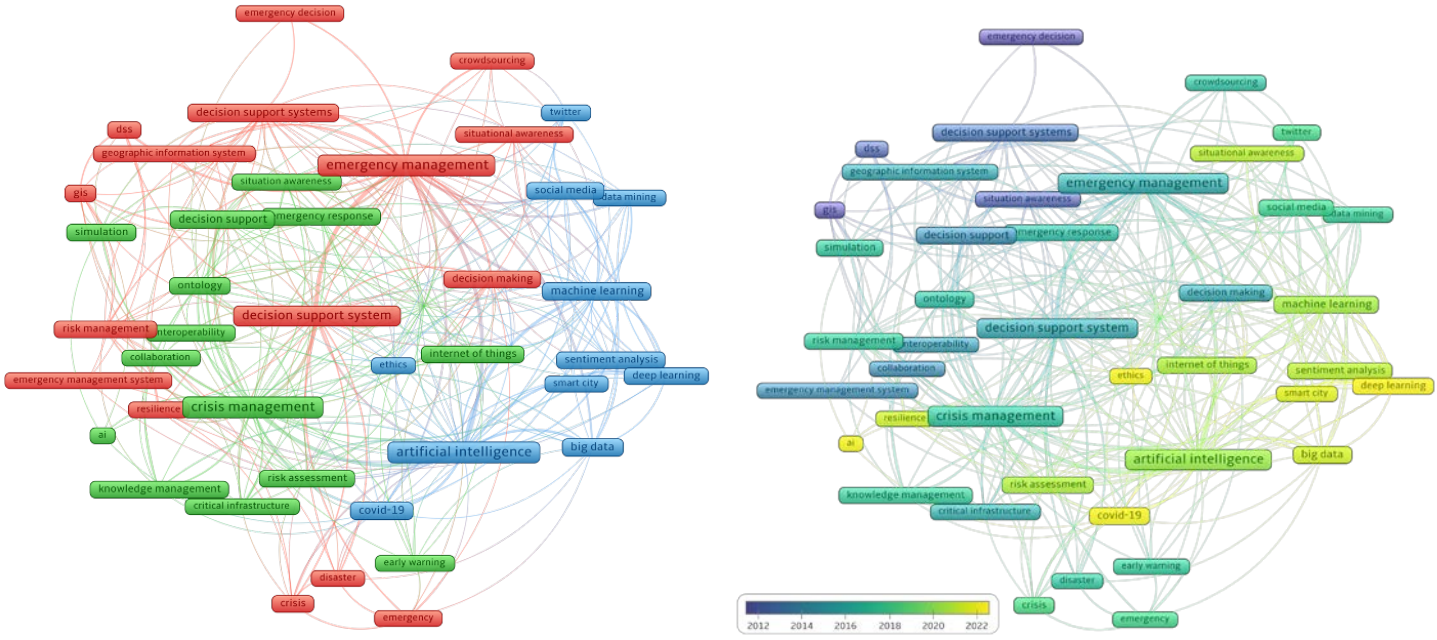


Figure 2. Usage of the authors keywords by clusters - on the right - of interest and year - on the left , search string on Scopus database *TITLE-ABS-KEY (("artificial intelligence" OR "AI") AND (emergency OR crisis) AND management AND survey)*. Generated with VOSviewer visualization software.

A keyword study (see Figure 2) highlights the use of AI for:

- **Achievements and purposes - red cluster** : Includes operational research and decision support systems to optimize decisions.
- **Knowledge-based systems - green cluster** : Includes systems for situational awareness or early warning systems based on expertise modeling.
- **Machine learning-based systems - blue cluster** : Includes big data (and especially social media data) processing with machine and deep learning methods.

The first cluster is excluded from this study, as it focuses on optimized decision-making rather than information processing itself.

In the following two sections, we provide an overview of the current state of the art in: (i) knowledge-based systems that enable static modeling of expertise on the basis of pre-defined concepts; (ii) systems based on learning approaches to solve specific problems in data processing; and (iii) systems that attempt to approach a multi-dimensional interpretation of data.

In order to target article searches in the Scopus database, the search strings were refined using keywords from the green (i) and blue (ii) clusters.

Knowledge-Based and Expert Systems : Symbolic Approaches

Knowledge-based and expert systems are based on a triptych consisting of a schema or structure for representing knowledge, a knowledge base subject to the representation schema, which embeds and computerizes the knowledge accumulated by domain experts, and a set of inference rules providing the system with logical reasoning for discovering or updating knowledge.

Numerous studies have focused on the implementation of such systems to support the information processing efforts of crisis decision-makers. Regarding knowledge representation, the terms meta-model (derived from model-driven engineering - MDE) and ontology are used as references: whatever the formalism used, it is a question of one way or another for experts to conceptualize the semantics of a domain, most often by qualifying the classes of interest

and the relationships that link them. Works such as (Alheadary, 2023; Christensen & Madsen, 2020; Correia et al., 2018; Kontopoulos et al., 2018) offer structured and possibly populated knowledge bases relating to the crisis management domain.

A reasoning layer, based on inference, can then be added on top of the knowledge base to take advantage of it, through the use of expert rules. Benaben et al. (2020), Rezaei and Vahidnia (2023), and Slam et al. (2015) propose this type of inference mechanism in their work, enabling them to offer users of the software developed an explicable framework of recommendations.

In order to make knowledge-based systems more reactive and autonomous, they have also been augmented by paradigms such as Complex Event Processing (CEP), enabling data to be received (from sensors, for example) and processed by rules in real time in order to relate it to knowledge already stored in the system, possibly triggering inferences. Systems of this kind, as proposed by Barthe-Delanoe¨ et al. (2014), Fertier et al. (2020), and Mijovic´ et al. (2019), have been widely used in the development of early warning systems.

In view of the uses to which they have been put, knowledge-based systems have undeniable advantages. In particular, they enable:

- Establish and unambiguously represent a semantic framework, thus ensuring unified interoperability;
- To a certain extent, to reason (by logical rules possibly integrating the notion of uncertainty) about knowledge with speed and explicability in order to draw conclusions.

Nevertheless, their lack of genericity and adaptability, due to a semantics fixed in a predefined semantic framework, makes it difficult for these systems to take advantage of heterogeneous data flows.

All these features qualify knowledge-based systems as ideal candidates for the realization of a COP.

Single-Task Performance Systems : Machine and Deep Learning Approaches

The use of AI learning techniques has been growing steadily over the past fifteen years. This growth can be relatively easily explained by the era of Big Data, with its increasing capacity to collect data (from connected and distributed sensors, imaging systems and social media, in particular) and store it at low cost, to drive machine learning and deep learning algorithms.

While the Scopus search "('crisis management' OR 'emergency management') AND ('machine learning' OR 'deep learning' OR 'big data' OR 'data mining')" has returned 792 papers from conferences and scientific journals since 2015 (as of February 13, 2024), adding the criteria "('social media' OR 'social network*' OR 'twitter' OR 'tweet*')" brings up 234 papers. In 2015, Imran et al. (2015) publish a survey on the use of AI methods to process social media messages for mass emergencies. Imran et al. explore three main challenges: data pre-processing, event detection and characterization, and aggregation into actionable information for end-users. Whatever the challenge addressed, the AI methods listed in this article remain classic, i.e. trained to perform the specific, often supervised tasks of topic modeling, message classification, named entity recognition and text summarization.

The observation can also be generalized to AI applications in crisis management for processing all other types of data (satellite and aerial imagery, sensors, remote sensing robotics, for instance), as shown by the very comprehensive systematic literature recently proposed by Sun et al. (2020) with the development of AIs highly specific to the tasks they perform. Kyrkou et al. (2023) raises the challenges that the state of the art is currently tackling:

- a broadening of the data types processed, using multi-modal models to solve problems of heterogeneity (joint processing of text and images, in particular)
- the ability to generalize and adapt AIs to new situations by generating learning data (through generative AI or the use of digital twins, lowering in the same way the cost and time of training datasets devlopment to supervise AIs)
- the enhancement of user confidence through explicability
- algorithm security

Today's learning-based AI methods are reviving the literature on automated data processing in crisis management with definite avatars:

- far greater generalization capacity than knowledge-based systems, making them highly adaptable
- deeper and deeper learning thanks to the use and fine-tuning of accessible foundation models, enabling excellent performance.

Nevertheless, work on the use of machine or deep learning algorithms in crisis management is focused on very specific tasks, the relevance of which is not always assessed in real time during the occurrence of a crisis whose stakes are always very specific (for example, different floods in different contexts for different responders require the analysis of aerial imagery in different ways). Furthermore, while the multimodal capability of algorithms now makes it possible to broaden the spectrum of types of data processed, the ability to select relevant data for processing and to produce different points of view and level of composing on the same system is not possible under the governance of these AIs.

All these features qualify ML and DL-based systems as ideal candidates for the realization of a SA.

Neuro-Symbolic Approaches

The SA paradigm suits fast individual decision-making, while the Sensemaking paradigm is for groups facing ambiguity, needing collective knowledge, inferences, and diverse perspectives to reach a shared understanding. The translation of this problem to AI is not obvious.

Although not new, neuro-symbolic AI approaches are currently undergoing a revival in tandem with advances in connectionist AI. The temptation is great to hybridize these two approaches in order to benefit from the advantages of each paradigm. Hitzler et al. (2022) notes a second reason for this flourishing, coming directly from the cognitive sciences: "we can understand artificial neural networks as an abstraction of the physical workings of the brain, while we can understand formal logic as an abstraction of what we perceive, through introspection". Wang et al. (2022) exposes the ability of human cognition to juggle 4 dimensions of symbolism and connectionism: Symbols vs Neurons, Deduction vs Induction, Compositionality vs Continuity and System 1 vs System 2. The first two dimensions are fairly common. It is nevertheless important to focus on the other two. *Compositionality vs Continuity Continuity* refers to the sequential activation of neurons in a continuous but variable signal (notably in the "weight" of information circulating) transmitted from neuron to neuron. *Compositionality* refers to the encoding of information in more or less large structures, enabling the generation of composite knowledge representations at variable levels of abstraction. *System 1 vs System 2* Kahneman (2011) proposes that human brain is composed of two conceptual systems. *System 1* is fast and intuitive but can be sometimes imprecise and *System 2* is logical, delibirative and conscious but slower and requires concentration.

In light of this, the literature shows that neuro-symbolic systems are particularly promising for implementing a form of dynamic between the perception of a situation and its understanding according to several levels of composition and expertise in dynamic environment under uncertainties, as required by the Sensemaking's *horizontal* and *vertical* processing of data. Recent articles such as (Booch et al., 2021; Dong et al., 2019) show efforts have been made in this direction. That said, Wang et al. (2022) notes that no hybridization paradigm currently allows us to effectively address the 4 dimensions between symbolism and connectionism sought in the Sensemaking process.

Limits of Current AIs to Meet Information Needs

In conclusion, several observations can be made concerning the state of the art. With regard to the information needs of decision-makers and the improvement of decision-support systems (particularly at this stage of understanding the situation), this is a complex and multifaceted cognitive and human process that is difficult to generalize. It involves coordinating different interpretations at different levels, depending on the scale of the crisis, the expertise involved and the often heterogeneous vocabularies, in a highly dynamic context that is unexpected at each subsequent event.

AI, with its advanced capabilities for processing massive data flows, has become a valuable ally in the development of modules for understanding crisis situations in real time in decision support systems. Symbolic AIs enable logical inferences to be defined and proposed on the basis of a vocabulary established by users, according to their needs. This includes early warning systems: for example, Fertier et al. (2020) offers an expert system based on listening to water level sensors combined with business rules collected from experts in the field (belonging to various institutions such as the Centre for studies and expertise on risks, the environment, mobility and development, Regional Department for the Environment, Planning and Housing, Regional Road Information and Coordination Centre, Interministerial zone headquarters). This type of system proves very effective to infer and confront information when the context of the crisis has already been anticipated and studied (even if it has not yet been encountered in the past), and is based on a strict vocabulary established in advance of the crisis. The dynamic context of the crisis is taken into account as

long as the business rules effectively respond to the conditions encountered in real time. Setting up such systems is laborious, not only in algorithmic terms but also time-consuming for the experts, given the complexity of the systems studied. This severely limits the adaptability and realism of such AIs, and shows the need to mix them with more flexible approaches.

ML and DL, by their mechanism of training by past experience, show remarkable capabities for adaptation and generalization, while drastically reducing the demands made on experts. Such systems are extremely useful for reducing the time taken by humans to monitor data, allowing them to devote their attention to tasks linked to the decision-making itself. For example, Suwaileh et al. (2022) offers such an AI dedicated to inferring geolocation of social media posts during a disaster for humanitarian purposes, based on very few data. Unlike an expert system, the AI allows extremely rapid inference, without having to draw up exhaustive lists of potential locations and their grammatical uses (an otherwise unrealistic task). However, these AIs respond to very specific tasks that need to be defined in advance of a crisis and cannot be created on the fly during a crisis, should new types of information prove interesting to extract.

While symbolic AIs make it possible to mimic expert control of the situation understanding process, ML and DL are AIs that are themselves controlled by humans, capable of carrying out tasks flexibly. So, with the aim of establishing AIs close to the needs of the Sensemaking process, neuro-symbolic approaches make it possible to imagine autonomous, flexible and realistic situation understanding modules in future decision support systems.

PROPOSAL

The finding shared by Kautz (2022) and Wang et al. (2022) lies in the promising potential of a neural architecture embedding symbolic reasoning engines, deemed closest to Systems 1 and 2 of Kahneman (2011), to support the Sensemaking process. Whether AI should follow the path of a biological or a non-biological architecture remains a large question. There are however strong beliefs that considering the notable progress of neurosciences, bio-inspired architecture are promising (Hole & Ahmad, 2021).

Neurosciences enable us to better understand these two systems, and most importantly now highlight a so-called inhibition System 3, which plays an executive role in regulating the use of the other two Systems (Houde′ & Borst, 2014). This key role makes System 3 the seat of human reasoning, in its ability to question preconceived perceptions and models, and to consider situations both holistically and with subtlety, especially thanks to the ability of meta-cognition given by System 3. However, while we have shown that the AI literature is interested in using the System 1 and System 2 paradigms to propose a hybridization architecture, no proposals have been made for System 3.

This section therefore aims to propose a new high-level integration architecture of a hybrid neuro-symbolic AI inspired by the System 3 allowing both *Compositionality* of the reasoning and *Continuity* of the information flow through neural networks.

We present this proposal in two parts. The first subsection focuses on the observations and theories currently proposed by the neurosciences around three biological subsystems - neurons, cortical columns and distributed inhibition processes - and the roles played by each in levels of human cognition. We use these 3 key biological subsystems to justify our proposal for a neuro-symbolic architecture, set out in the second subsection.

Three Key Biological Sub-Systems for the Human Cognition

In his book *A thousand brains: A new theory of intelligence*, Hawkins breaks down human cognitive capacity into three sub-parts that are found in the human neocortex possessing both *hierarchical - i.e. vertical, in columns -* and *distributed - i.e. horizontal, interaction and voting -* roles.

Neurons

Neurons are the basic building blocks of the human neocortex, transmitting information through impulses. When stimulated, a neuron generates an action potential that travels down its axon, leading to the release of neurotransmitters at the synapse, facilitating communication with other neurons. The structure of neural networks is dynamically influenced by experiences through a phenomenon known as neuroplasticity. When engaging in new situations, the connections between neurons, can be strengthened or weakened. Over time, the observations show that repeated experiences can lead to long-lasting changes in neural pathways, contributing to the continuous evolution of the brain's structure and functionality.

Thanks to their ability to transmit the information in a distributed and loose way, neurons ensure the *Continuity* of the cognitive system. Moreover, the neural map at time t represents the knowledge stored by the individual at this exact instant.

Cortical Columns : One Model Generation

The cortical columns of the neocortex are highlighted by Mountcastle's work in 1957 (Mountcastle, 1957). Cortical columns are sensory-motor systems of the neocortex made up of vertically aligned neurons, all with the same laminar structure and consisting of "functional units of information processing". They perceive data from outside, process it through a back and forth circulation of data and provide specific models of sub-systems of the environment. According to Hawkins et al. (2017), cortical columns "learn complete predictive models of observed objects". Thanks to neural sparse connections between columns, one cortical column can infer models of objects based on the partial knowledge of the adjacent columns, hence ensuring low-level information complementarity in the construction of a model.

The ability of the columns to create atomic models of objects in the environment from individual points of view contributes to the cognitive capacity of *Compositionality*. Models can be closely related to symbolic representations of the environement. Moreover, the ability to inhibit the flow of information from neuron to neuron within the columns at the neuronal level, also associates it with capacities close to those of Kahneman's two systems.

Inhibition and Voting Process : Several Models to One Perception

Mountcastle (1997) observes that columns can also be linked to each other by long-range connections. Hawkins (2021) defends the existence of a voting process by which the thousands of models produced by cortical columns offering different points of view of the same environment can be reconciled into a single, also stable, perception of the environment. These long-range connections enable precisely this process of voting between columns.

Houde´ and Borst (2014) describe a comparable inhibition process that associates Kahneman's two systems with a higher-level executive system (meta-cognitive) that is nonetheless clearly distributed within the neocortex. Inhibition System 3 enables us to implement a conscious deductive approach in response to unknown or complex situations.

The regulatory capacity provided by these two complementary theories enables a high-level compromise between *Compositionality Continuity* by reconciling the model of an object underpinning several points of view and levels of submodels, and thereby ensuring cognitive consistency over time. Such higher-level inhibitory system (at a meta-cognition level) is typically not proposed by Hole and Ahmad (2021).

Perspectives for a Bio-Inspired Neuro-Symbolic AI Architecture for Sensemaking Support

Inter-Disciplinary Framework

Figure 3 presents a framework that summarizes and justifies how neurosciences can, through the three levels of neurons, cortical columns and inhibition and voting processes, inspire the creation of AI responding to the very specific informational needs questions demanded by the three paradigms of COP, SA and Sensemaking.

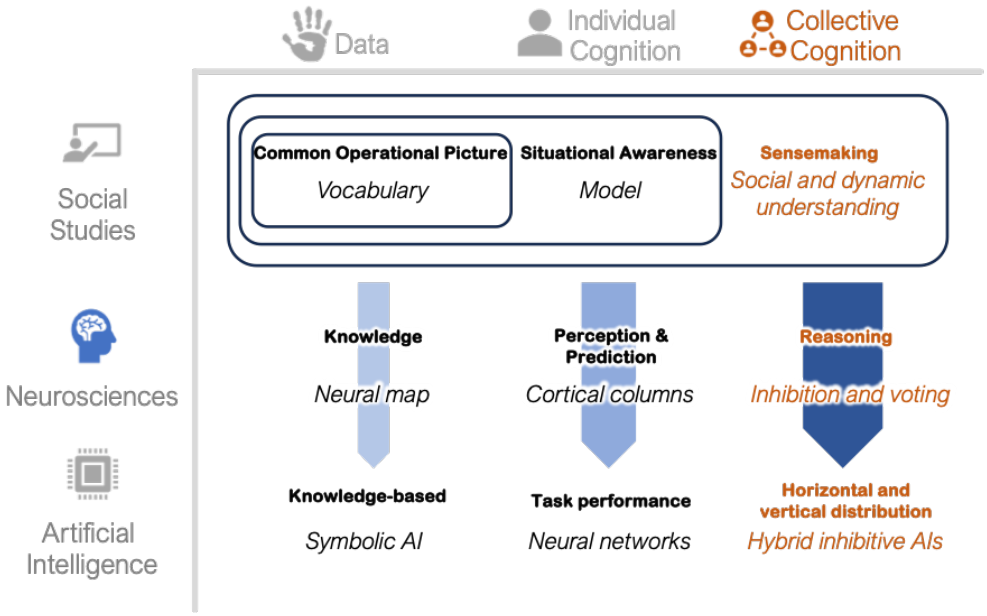


Figure 3. Neuro-sciences as a promising pivot to create new AI aligned with the crisis managers' informational needs.

Bio-Inspired Neuro-Symbolic AI Architecture

Based on neuroscientific observations of human reasoning, we can extract a number of specifications for the construction of a bio-inspired neuro-symbolic AI, capable of embodying the notions of *Compositionality*, *Continuity* and those contained in Kahneman, Houde´ and Borst's *Three Systems* and support Sensemaking.

- **Sequencing neural networks within columns**: sequencing neural networks ensures *Continuity* in the flow of information. On the one hand, deeper neural networks enable a complex level of abstraction, by producing more higher-level of aggregation of the input data. In terms of sense Sequencing neural networks thus addresses the *verticality* problem of Sensemaking.
- **Parallelization of columns**: in the same way that cortical columns are used in human cognition as individual computational units enabling atomic models (a point of view, a subsystem - *Compositionality*) of an environment to be established in a parallelized processing process with different data, we propose to parallelize neural networks. The more neural networks there are in parallel, the wider the scope of possible perception by integrating more data sources. In order to share partial models, like adjacent cortical columns, neural connections can be set up between columns. Parallelization of neural networks thus makes it possible to address the *horizontality* problem outlined of Sensemaking.
- **Executive system distributed on two levels**: an executive inhibition system regulates the use of columns on two levels. At a meta-level, a voting process is used to unify the models produced by the columns into a situation model delivered to the crisis manager, thus meeting contextualized information needs. To make this possible, a second process, this time inhibitive, is implemented to regulate the use of columns according to their relevance (i.e. signals perceived from the outside) and the connections between columns. The interest of such an executive system is to ensure coherence between (i) the informational models generated and (ii) the cognitive reasoning implemented within and between the columns. Thus, this executive system is at the heart of the concerns in the development of such an AI architecture: it no longer addresses vertical and horizontal problems independently, but proposes an architecture that is adaptive to needs on both dimensions.

CONTRIBUTIONS, LIMITS AND PERSPECTIVES

This paper makes two contributions: (i) while no AI in the literature is currently capable of supporting the collective cognitive effort made during Sensemaking, current knowledge of the cognitive mechanisms of the human brain observed in neurosciences allows us to propose a pivot between informational needs and the implementation of adapted AIs ; (ii) the proposal of a new AI architecture based on voting and inhibition mechanisms distributed between cortical columns enables us to envisage meta-cognitive capacities capable of supporting Sensemaking in its *horizontal* and *vertical* dimensions (see contributions on Figure 3).

That being said, the proposal remains at its early stage and especially needs further investigations that are the current perspectives of this proposal:

- **Methodologies to better understand mechanisms involved in human cognitive executive system** : current medical imaging techniques do not allow to observe inhibition behaviours at the neuronal level. Hence, very specific experimental methodologies must be implemented to allow a proper control of all parameters and to enable in-depth conclusions to be drawn.
- **Transferring human tests to AI tests**: tests usually carried out by humans to observe their cognitive mechanisms are not directly transferable to AI (for example, the Strooper test - Color-Word Interference Task - requires inhibition on the part of the human, but remains a very simple test for computers).
- **Evaluation of the AI performances** : new measures need to be created to (i) evaluate the predictions generated by this type of AI, given that they would not be dedicated to a single task (ii) and evaluate the added value of this AI according the two-dimensions (horizontal and *vertical*) of Sensemaking.

ACKNOWLEDGEMENT

We acknowledge the use of an AI for translation purposes.

REFERENCES

Alheadary, W. G. (2023). Artificial intelligence metamodel for controlling and structuring the crisis management domain. *Journal of Computer Science*, *19*(8), 988–997. https://doi.org/10.3844/jcssp.2023.988.997

Barthe-Delanoe¨, A.-M., Truptil, S., Be´naben, F., & Pingaud, H. (2014). Event-driven agility of interoperability during the run-time of collaborative processes. *Decision Support Systems*, *59*, 171–179.

Benaben, F., Fertier, A., Montarnal, A., Mu, W., Jiang, Z., Truptil, S., Barthe-Delanoe¨, A.-M., Lauras, M., Mace-Ramete, G., Wang, T., et al. (2020). An ai framework and a metamodel for collaborative situations: Application to crisis management contexts. *Journal of Contingencies and Crisis Management*, *28*(3), 291–306.

Benaben, F., Montarnal, A., Fertier, A., & Truptil, S. (2016). Big-data and the question of horizontal and vertical intelligence: A discussion on disaster management. *Collaboration in a Hyperconnected World: 17th IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2016, Porto, Portugal, October 3-5, 2016, Proceedings 17*, 156–162.

Booch, G., Fabiano, F., Horesh, L., Kate, K., Lenchner, J., Linck, N., Loreggia, A., Murgesan, K., Mattei, N., Rossi, F., et al. (2021). Thinking fast and slow in ai. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(17), 15042–15046.

Brown, A. D., Colville, I., & Pye, A. (2015). Making sense of sensemaking in organization studies. *Organization studies*, *36*(2), 265–277.

Christensen, L. L. W., & Madsen, B. N. (2020). A danish terminological ontology of incident management in the field of disaster management. *Journal of Contingencies and Crisis Management*, *28*(4), 466–478.

Coche, J., Kropczynski, J., Montarnal, A., Tapia, A., & Benaben, F. (2021). Actionability in a situation awareness world: Implications for social media processing system design. *ISCRAM 2021-18th International conference on Information Systems for Crisis Response and Management*, (2391), p–994.

Correia, A., Severino, I., Nunes, I. L., & Simo˜es-Marques, M. (2018). Knowledge management in the development of an intelligent system to support emergency response. *Advances in Human Factors and Systems Interaction: Proceedings of the AHFE 2017 International Conference on Human Factors and Systems Interaction, July 17- 21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8*, 109–120.

Dong, H., Mao, J., Lin, T., Wang, C., Li, L., & Zhou, D. (2019). Neural logic machines. *arXiv preprint arXiv:1904.11694*.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors*, *37*(1), 32–64.

Fertier, A., Barthe-Delanoe¨, A.-M., Montarnal, A., Truptil, S., & Be´naben, F. (2020). A new emergency decision support system: The automatic interpretation and contextualisation of events to model a crisis situation in real-time. *Decision Support Systems*, *133*, 113260.

Flach, J. M. (2015). Situation awareness: Context matters! a commentary on endsley. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 59–72.

Hawkins, J. (2021). *A thousand brains: A new theory of intelligence*. Basic Books.

Hawkins, J., Ahmad, S., & Cui, Y. (2017). A theory of how columns in the neocortex enable learning the structure of the world. *Frontiers in Neural Circuits*, *11*. https://doi.org/10.3389/fncir.2017.00081

Hitzler, P., Eberhart, A., Ebrahimi, M., Sarker, M. K., & Zhou, L. (2022). Neuro-symbolic approaches in artificial intelligence. *National Science Review*, *9*(6), nwac035.

Hole, K. J., & Ahmad, S. (2021). A thousand brains: Toward biologically constrained ai. *SN Applied Sciences*, *3*(8), 743.

Houde´, O., & Borst, G. (2014). Measuring inhibitory control in children and adults: Brain imaging and mental chronometry. *Frontiers in psychology*, *5*, 616.

Imran, M., Castillo, C., Diaz, F., & Vieweg, S. (2015). Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)*, *47*(4), 1–38.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus; Giroux.

Kautz, H. (2022). The third ai summer: Aaai robert s. engelmore memorial lecture. *AI Magazine*, *43*(1), 105–125.

Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. *Expertise out of context: Proceedings of the sixth international conference on naturalistic decision making*, *113*.

Kontopoulos, E., Mitzias, P., Moßgraber, J., Hertweck, P., van der Schaaf, H., Hilbring, D., Lombardo, F., Norbiato, D., Ferri, M., Karakostas, A., et al. (2018). Ontology-based representation of crisis management procedures for climate events. *ISCRAM*.

Kuusisto, R., Kuusisto, T., & Armistead, L. (2005). Common operational picture, situation awareness and information operations. *ECIW*, 175–186.

Kyrkou, C., Kolios, P., Theocharides, T., & Polycarpou, M. (2023). Machine learning for emergency management: A survey and future outlook. *Proceedings of the IEEE*, *111*(1), 19–41. https://doi.org/10.1109/JPROC.2022.3223186

Leighton, J. P., & Sternberg, R. J. (2004). *The nature of reasoning*. Cambridge University Press.

Mijovic´, V., Tomasˇevic´, N., Janev, V., Stanojevic´, M., & Vranesˇ, S. (2019). Emergency management in critical infrastructures: A complex-event-processing paradigm. *Journal of Systems Science and Systems Engineering*, *28*, 37–62.

Mountcastle, V. B. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of neurophysiology*, *20*(4), 408–434.

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain: a journal of neurology*, *120*(4), 701–722.

Muhren, W. J., & Van de Walle, B. (2010). A call for sensemaking support systems in crisis management. In *Interactive collaborative information systems* (pp. 425–452). Springer.

Palen, L., Anderson, J., Bica, M., Castillos, C., Crowley, J., D´ıaz, P., Finn, M., Grace, R., Hughes, A., Imran, M., et al. (2020). Crisis informatics: Human-centered research on tech & crises.

Palen, L., & Anderson, K. M. (2016). Crisis informatics—new data for extraordinary times. *Science*, *353*(6296), 224–225.

Pirolli, P., & Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of international conference on intelligence analysis*, *5*, 2–4.

Reuter, C., Hughes, A. L., & Kaufhold, M.-A. (2018). Social media in crisis management: An evaluation and analysis of crisis informatics research. *International Journal of Human–Computer Interaction*, *34*(4), 280–294.

Rezaei, Z., & Vahidnia, M. H. (2023). Effective medical center finding during covid-19 pandemic using a spatial dss centered on ontology engineering. *GeoJournal*, *88*(3), 2721–2735.

Slam, N., Wang, W., Xue, G., & Wang, P. (2015). A framework with reasoning capabilities for crisis response decision–support systems. *Engineering Applications of Artificial Intelligence*, *46*, 346–353.

Sun, W., Bocchini, P., & Davison, B. D. (2020). Applications of artificial intelligence for disaster management. *Natural Hazards*, *103*(3), 2631–2689.

Suwaileh, R., Elsayed, T., Imran, M., & Sajjad, H. (2022). When a disaster happens, we are ready: Location mention recognition from crisis tweets. *International Journal of Disaster Risk Reduction*, *78*, 103107.

Wang, W., Yang, Y., & Wu, F. (2022). Towards data-and knowledge-driven artificial intelligence: A survey on neuro-symbolic computing. *arXiv preprint arXiv:2210.15889*.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization science*, *16*(4), 409–421.

Zade, H., Shah, K., Rangarajan, V., Kshirsagar, P., Imran, M., & Starbird, K. (2018). From situational awareness to actionability: Towards improving the utility of social media data for crisis response. *Proceedings of the ACM on human-computer interaction*, *2*(CSCW), 1–18.

Envisioning the Future of Emergency Management focused on Artificial Intelligence

Ana-Gabriela Núñez
Department of Computer Science,
Universidad de Cuenca
anag.nunez@ucuenca.edu.ec

Patricia Quiroz-Palma
Universidad Laica Eloy Alfaro de Manab´ı
patricia.quiroz@uleam.edu.ec

Ma Carmen Penadés
IUMTI - Universitat Politècnica de València,
Spain
mpenades@dsic.upv.es

Otto Parra
Department of Computer Science,
Universidad de Cuenca
otto.parra@ucuenca.edu.ec

Oscar Alvear
Department of Computer Science,
Universidad de Cuenca
oscar.alvear@ucuenca.edu.ec

ABSTRACT

Artificial Intelligence (AI) in Emergency Management integration has become crucial to improving crisis response efficiency and effectiveness. AI presents new paradigms for responding to emergencies but poses technical and operational challenges, ethical considerations, and the need for robust regulatory frameworks. The study examines various AI applications in emergency scenarios and employs a systematic literature review methodology to identify solutions and future directions for ethical and effective AI integration. The findings highlight the critical balance between leveraging AI capabilities and ensuring safety, privacy, and ethical compliance. Understanding AI's full scope and potential in this critical field is essential as we navigate technological advancements. The article contributes to understanding AI's potential to transform emergency management practices and emphasizes the importance of ethical considerations and regulatory frameworks.

Keywords

Artificial Intelligence, Emergency Management, Ethical Considerations, Technologies.

INTRODUCTION

Integrating Information and Communication Technologies (ICT) and Artificial Intelligence (AI) in emergency management is a transformative development representing a significant shift in how societies respond to crises. The deployment of AI applications in emergency scenarios has the potential for profound positive impacts but also carries inherent risks, thus necessitating a rigorous ethical framework. The central focus of this discourse is the application of AI to enhance the efficacy of emergency responses while navigating the complex ethical landscapes these technologies inhabit.

AI-powered tools in emergency management have the potential to revolutionize the traditional approaches to crisis response. They leverage advanced signal processing and systems analysis techniques commonly used in fields such as aeronautical engineering and medical diagnostics. When applied to emergency management, these methodologies

provide unprecedented capabilities in real-time evaluation of complex and dynamic situations, offering actionable intelligence to support informed decision-making.

The widespread use of sensors has resulted in the generation of rich data streams, which provide a unique opportunity to enhance emergency response situational awareness and responsiveness. Additionally, advancements in AI methodologies, such as Machine Learning (ML) and cognitive computing, offer potential solutions for addressing the complex challenges of large-scale emergencies (Freeman, 2020; Yadav, 2024). Emergency management can significantly improve efficiency and effectiveness by utilizing AI for predictive analysis, incident detection, resource allocation, and deployment. These developments offer the possibility of reducing human and financial losses and promoting more sustainable environmental practices, which can mitigate the pressing global challenges through innovative AI applications.

The main aim of this section is to carry out a targeted literature review to highlight the recent developments and emerging trends in integrating AI into emergency management. The study is centred around three key areas: firstly, exploring how AI tools can improve the efficiency and effectiveness of emergency management while identifying any associated technical and operational challenges and secondly, examining the necessary regulatory frameworks for the safe use of AI in emergency management. Lastly, the study delves into the ethical implications of employing AI in emergency management.

The structure of this paper is outlined as follows: Section II discusses the background, highlighting the role of AI in emergency management, including technological challenges and ethical considerations. Section III describes the methodology used for the systematic literature review. Section IV analyzes and presents findings from the literature review. Finally, the paper concludes with a discussion and conclusions, emphasizing the potential for future research, the strengths and weaknesses identified in the study, and the overall conclusion.

BACKGROUND

AI is undergoing rapid change, particularly in its application to emergency management. However, this integration of AI poses both technological challenges and ethical considerations throughout all phases of the emergency management cycle, from preparation to response and post-disaster recovery. Ethical considerations must be prioritized when deploying AI technologies in emergency situations, ensuring adherence to ethical standards and promoting equitable outcomes. Furthermore, AI's role in promoting sustainable development and enhancing resilience is crucial, highlighting the broader implications of AI applications beyond immediate emergency response scenarios (Freeman, 2020).

AI will play a critical role in emergency management by enhancing the abilities of first responders to handle various crises, including natural disasters and industrial incidents. This is achieved by developing advanced Situation Awareness (SA) models that incorporate real-time sensor data, enabling dynamic decision-making. The European ASSISTANCE project is a prime example of this, as it has created an actionable SA (aSA) module that integrates AI technology with sensor data to predict and display gas distributions in emergency scenarios. This module aids first responders in their decision-making processes. (Mioch et al., 2021)

A study conducted in Xuwen County, China, examined the impact of AI applications on public perception and trust related to emergency management services. The study found a significant positive correlation between public expectations of emergency management services and their satisfaction with these services. The introduction of AI governance, characterized by scientific prediction, prevention, and management, plays a vital role in mitigating the damage caused by natural disasters and strengthens public trust in government services. This shift towards AI-driven governance can revolutionize traditional emergency management models, fostering a more harmonious relationship between the government and the public (Chen et al., 2023).

The findings suggest that in the future, emergency management will heavily depend on AI technologies for efficient response and to satisfy the public. By integrating AI in emergency response, decision-making processes can be streamlined, and a trust-based relationship can be established between emergency management authorities and the public. This advancement implies a more responsive, efficient, and publicly endorsed emergency management framework that harnesses AI's capabilities to meet the complex demands of modern crises.

RESEARCH METHOD

A systematic review uses a detailed, reliable, iterative, and thorough methodology, which includes planning, conducting, and reporting the review, as outlined by the Kitchenham and Charters, 2007 methodology.

Planning the Review

This systematic review aims to gather and analyze the tendencies, challenges, and solutions to provide a nuanced theoretical understanding and practical strategies for the ethical integration of AI technologies into emergency management systems. This work will support a discourse on ensuring that AI implementation aligns with ethical standards and supports the resilience and efficacy of emergency responses.

Three pivotal research questions guide this study, each reflecting the urgent need to explore the expansive role of AI within the domain of emergency management:

RQ1. How can AI tools enhance emergency management, and what are the inherent technical and operational challenges?

RQ2. What regulatory framework is required to support AI's safe and ethical adoption in emergency management while fostering innovation?

RQ3. What are the ethical implications of employing AI in emergency management, and how can we ensure transparency and accountability in its applications?

Data Source Identification and Search Strategy

To collect primary studies, we referred to the digital libraries of prominent organizations, such as Scopus, WoS, and the Iscram Library. Our search focused on the past five years, from 2019 to January 2024, to capture the latest trends in AI.

Our study covered conference proceedings and journal articles, thoroughly examining paper titles and abstracts. Considering their unique syntax, it's important to mention that the search string notation was tailored for each digital library. The search query employed was:

- Scopus and WoS : ("Artificial Intelligence" OR "AI") AND ("emergency management" OR "crisis management" OR "disaster response") AND ("regulations" OR "ethical considerations" OR "legislation" OR "law*") AND ("tools" OR "technologies" OR "applications")
- Iscram Library: ("Artificial Intelligence" OR "AI")

Criteria for Selecting Primary Studies

The criteria for selecting primary studies involve an automated search through digital libraries using a predefined search string. However, it's possible for retrieved articles to not directly align with the study's focus, even if they contain the search string terms in their titles and abstracts. Hence, to ensure a consistent approach in including or excluding articles, we refer to the criteria outlined in Table 1. Articles meeting these criteria will be accepted or rejected based on their alignment with the study's focus.

Quality Assessment

In assessing studies selected for our investigation on the integration of AI in emergency management, we applied predefined inclusion, exclusion, and quality assessment criteria. Initially, we identified studies that specifically addressed the application of AI in emergency management scenarios, including those that discussed technical and operational challenges, regulatory frameworks, and ethical implications and provide case studies or practical applications. Studies not focused on AI in emergency management, lacking in academic rigour, or not within the scope of our research questions were excluded. For the quality assessment, each study underwent a thorough evaluation by five reviewers, focusing on criteria such as the clarity of AI application, the depth of analysis regarding challenges, discussions on regulatory and ethical considerations, methodological rigour, relevance and currency of the research, and its contribution to the field. Scores were assigned as Low, Medium, or High based on a consensus between the reviewers for each criterion. This rigorous process ensured that only studies of high relevance and quality were considered for our analysis, facilitating a robust synthesis of current knowledge and practices in using AI for emergency management.

Table 1. Inclusion and Exclusion Criteria for AI in Emergency Management Research

Criteria Type	Inclusion Description	Exclusion Description
Thematic Scope	Focuses on AI applications in emergency management, including technical and operational challenges.	Irrelevant to AI or emergency management, covering areas not focused on these topics.
Regulatory and Ethical Framework	Examines regulatory and ethical frameworks for safe and ethical AI adoption in emergency management.	Works that do not address regulatory or ethical aspects of employing AI in emergency scenarios.
Publication Type	Peer-reviewed journal articles and conference reports providing innovative insights.	Non-peer-reviewed publications, blogs, news articles, and non-academic content.
Publication Period	Published within the last five years for relevance.	Publications older than five years, considered technologically obsolete.
Language	English publications for international accessibility.	Non-English documents without available translation.
Research Methodology	Empirical, qualitative, quantitative, or mixed methods studies.	Opinions, editorials, and articles without a solid empirical or theoretical foundation.
Case Studies and Practical Applications	Includes real-world applications or case studies of AI in emergencies.	Theoretical studies without practical applications or real-world case studies.

Data Extraction Strategy

We categorized the research inquiries into various criteria to extract data from the chosen primary studies. Table 3 provides an overview of the data extraction criteria (EC) for classifying the gathered solutions and applications. The data extraction criteria and their options are described as follows:

EC1 *Innovations and Technologies*: What innovations and technologies are being utilized or explored in this context?

EC2 *Domain Described*: What is the primary domain or area of study where these AI applications are being utilized?

EC3 *Technical and Operational Challenges*: What technical and operational challenges are faced when implementing these AI applications?

EC4 *Impact on Efficiency and Effectiveness*: How do these AI applications impact the efficiency and effectiveness of the operations?

EC5 *Case Studies or Practical Application Examples*: Are there any case studies or practical examples where AI has been applied successfully in this field?

EC6 *Existing Regulatory Frameworks*: What regulatory frameworks currently govern the use of AI in this domain?

EC7 *Ethical Considerations*: What ethical considerations must be considered with the use of AI in this area?

EC8 *Model Used*: What AI models or methodologies are being employed in these applications?

The criteria corresponding to EC1, EC2, EC3, EC4, and EC5 correspond to RQ1, EC6 to RQ2, and EC7 and EC8 to RQ3.

Conducting the Review

The process of identifying primary studies in digital libraries took place in January 2024, resulting in the acquisition of thirty-five papers. There were instances of duplication among these papers due to their presence in multiple digital libraries. Table 2 shows the distribution of papers across different digital libraries, along with the screening process at each stage, culminating in the total number of studies incorporated into the review.

The selection of papers followed specific inclusion and exclusion criteria. Notably, many papers sourced from digital libraries were initially considered potential candidates based on the presence of the search string in their titles and abstracts. However, upon closer examination, including full paper scans, it became evident that the topics diverged from the scope of our investigation. Consequently, eighteen papers were ultimately selected for our study. The selected studies by the database are shown in Figure 1.

Table 2. Results of conducting stage

Source	*Potential Studies*	*Removing duplicates*	*Scanning title and abstract*	*Selected Studies*
Scopus	16	16	12	6
WoS	6	2	1	0
Iscram Library	13	13	12	12
Total	35	31	25	**18**

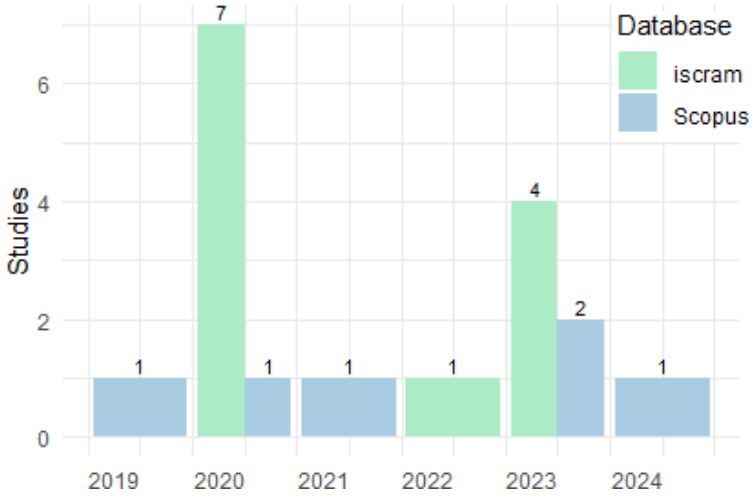


Figure 1. Annual Distribution of Selected Studies by Database

ANALYSIS AND FINDINGS FROM THE LITERATURE REVIEW

The systematic review found eighteen selected studies. Table 3 summarises the extracted information quantitatively (for a detailed synthesis of the selected studies, please refer to Appendix A).

Each data extraction criterion and the obtained results are discussed.

EC1. Innovations and Technologies

In the domain of emergency management, recent literature has illuminated the significant role of AI and ML in pioneering Innovations and Emerging Technologies. Machine Learning is the most prominent area, accounting for 27.78% of the studies (Farasin et al., 2020; Imran et al., 2020; Ortiz et al., 2020; Palomba et al., 2020; Ramsey et al., 2023). This is followed by Generative AI (Freeman, 2020; Karinshak, 2024; Tsai et al., 2023) and Neural Networks (Jiang et al., 2021; Moumtzidou et al., 2020; Wu & Shan, 2021), each comprising 16.67% of the research. Decision Support Systems, although significant, represent a smaller portion at 11.11% (Nussbaumer et al., 2023; Suciu et al., 2021). These results demonstrate a diverse range of interests within the field of AI, highlighting Machine Learning as the most extensively explored area.

Systems like Sahana EDEN exemplify the integration of AI into disaster management, enhancing emergency response capabilities (Nussbaumer et al., 2023). The serious game format has been utilized to elucidate AI applications in crisis contexts (Adam & Lauradoux, 2022), while the synergy of crowdsensing with AI signifies a novel approach to emergency solutions (El Barachi et al., 2020). The emergence of commercial gaming technologies such as Oculus Rift and Photon Cloud marks a notable advancement in creating immersive environments for training emergency responders (Sharma et al., 2019). The CitizenHelper system (Pandey et al., 2020) and the adoption of platforms like Google Dialogflow for chatbots (Tsai et al., 2023) demonstrate the innovative use of real-time, AI-infused analytics and conversational interfaces in emergency management training. Further, computational creativity has been applied to scenario modelling, and pyramid ConvLSTMs have been introduced to enhance high-dimensional data analysis for emergency scenarios (De Nicola et al., 2019; Jiang et al., 2021). Integrating VR for immersive training and AI for data analysis reflects a commitment to developing sophisticated training solutions (Xia et al., 2022). Progress in flood detection has been achieved through semantic segmentation and

ML, while AI's predictive analytics are increasingly utilized for emergency preparedness and response (Ortiz et al., 2020; Suciu et al., 2021). Combining AI with crowdsourced human feedback for image categorization (Imran et al., 2020). Applying Random Forest and U-Net models to satellite imagery is a testament to the evolving intersection of human intelligence and AI (Palomba et al., 2020). The innovative application of deep learning to satellite data, especially using Sentinel-2's capabilities, transforms crisis management (Farasin et al., 2020). Moreover, developing interactive visualizations for explaining complex ML models enhances the interpretability of AI (Ramsey et al., 2023). AI-based predictive dispatch for emergency vehicle routing underscores the efficiency gains possible in emergency fleet management (Landsberg et al., 2020). Collectively, these studies showcase a trend towards technologically sophisticated, efficient, and adaptive emergency management systems.

EC2. Domain Described

In the context of emergency management and crisis management, studies have explored a diverse range of areas. This theme represents the primary focus, accounting for 83.33% of 15 studies. The research encompasses natural disaster scenarios (Karinshak, 2024; Moumtzidou et al., 2020; Nussbaumer et al., 2023; Pandey et al., 2020; Tsai et al., 2023; Xia et al., 2022), education and society (Adam & Lauradoux, 2022), and the use of crowdsensing and social media (El Barachi et al., 2020). Integrating satellite-based flood mapping in crisis management is another significant area, as shown in Moumtzidou et al., 2020.

Public health, another important domain, contributes 16.67% with three studies (Adam & Lauradoux, 2022; Ramsey et al., 2023; Wu & Shan, 2021). Training, essential for AI application, comprises 11.11% (Pandey et al., 2020; Xia et al., 2022). Among these, studies Wu and Shan, 2021 and Sharma et al., 2019 delve into social governance and public health emergencies, underscoring AI's crucial role in complex scenarios. The deployment of technologies in smart cities for pandemic management and earthquake risk is detailed in (Pandey et al., 2020) and (Tsai et al., 2023).

EC3. Technical and Operational Challenges

Technical and Operational Challenges constituted 94.44% of the research with 17 studies. This is followed by Emergency Management Challenges, which involve seven studies and account for 38.89% of the total research. Public Health Challenges and Disaster Response Challenges, each covered by 1 study (5.56%), represent a smaller portion of the research. Research addressing emergency management's technical and operational challenges highlights this field's diverse and complex hurdles. The unpredictable and variable nature of disasters poses significant challenges, demanding the development of robust and resilient response mechanisms (Nussbaumer et al., 2023). In serious gaming for disaster education, the primary challenge is creating engaging and informative experiences with operational relevance (Adam & Lauradoux, 2022). Integrating crowdsourced data into AI algorithms for emergency response faces the critical need for accuracy and timeliness in data analysis (El Barachi et al., 2020).

The paper Sharma et al., 2019 examines the operational challenges in emergency management, including the complex orchestration of multiple systems and stakeholders. The convergence of Virtual Reality (VR) and Artificial Intelligence (AI) in creating immersive training environments, along with the intricacies of data collection and processing, is emphasized in Xia et al., 2022.

Managing and effectively auguring vast datasets to enhance disaster response is a core focus, as delineated in (Moumtzidou et al., 2020; Wu & Shan, 2021), while Ortiz et al., 2020 concentrates on integrating varied data sources for dependable emergency management. Data preprocessing and model training, particularly challenging with limited data during emergencies, are significant operational hurdles, as discussed in Palomba et al., 2020. Additionally, Suciu et al., 2021 explores the complexities of merging different data sources and ensuring interoperability amidst a pandemic. The delicate balance required between AI precision and human judgment in damage assessment tasks is critically explored in Imran et al., 2020.

EC4. Impact on Efficiency and Effectiveness

Research emphasizes enhancing emergency response capabilities through AI-driven real-time data analysis and improving efficiency in predictive tasks (Adam & Lauradoux, 2022; El Barachi et al., 2020; Farasin et al., 2020; Landsberg et al., 2020; Suciu et al., 2021; Wu & Shan, 2021). Sharma et al., 2019 showcases the effective deployment of AI agents in collaborative virtual environments designed for emergency response training. Ramsey et al., 2023 examines the role of visualizations in enhancing the interpretation of model decisions to boost operational efficiency. Moreover, Landsberg et al., 2020 explores the potential of AI to refine dispatch strategies, which could lead to faster response times and improved patient outcomes.

Studies investigating the operational impacts highlight resource allocation challenges and the vulnerabilities arising from increased dependence on technology (Jiang et al., 2021; Nussbaumer et al., 2023; Pandey et al., 2020; Tsai

Table 3. Summary of the Result (N=18)

Extraction Criteria (EC)	Criteria	# Studies	%
EC1	Specific AI Applications	18	100%
	Decision support systems	2	11.11%
	Serious game	1	5.56%
	Crowdsensing	1	5.56%
	AI Simulation	2	11.11%
	Machine Learning	5	27.78%
	Generative AI	3	16.67%
	Virtual Reality	1	5.56%
	Neural Networks	3	16.67%
EC2	Emergency/Crisis Management	15	83.33%
	Public Health	3	16.67%
	Training	2	11.11%
EC3	Technical and Operational Challenges	17	94.44%
	Emergency Management Challenges	7	38.89%
	Public Health Challenges	1	5.56%
	Disaster Response Challenges	1	5.56%
EC4	Impact on Efficiency and Effectiveness	18	100%
	Real-time data analysis	6	33.33%
	Decision-making during response	5	27.78%
	Technological Reliance	4	22.22%
EC5	General mentions of case studies or practical application	15	83.33%
	Without specific details provided	3	16.67%
EC6	Adherence to Data Protection Laws	3	16.67%
	Alignment with Privacy Regulations	7	38.89%
	Focus on technical aspects	6	33.33%
	Not specified	2	11.11%
EC7	Ethical Considerations	18	100%
	Privacity	12	66.67%
	Bias	3	16.67%
	Data	9	50%
EC8	Model Used	14	77.78%
	Not specified	4	22.22%

et al., 2023). Selected studies also illuminate AI's advanced understanding and decision-making capabilities during crises (Adam & Lauradoux, 2022; Jiang et al., 2021). Additionally, various studies collectively identify significant improvements in training methods, flood detection, and the overall efficiency of emergency response (Moumtzidou et al., 2020; Ortiz et al., 2020; Xia et al., 2022)

EC5. Case Studies or Practical Application Examples

The results strongly focus on practical applications and case studies in AI and emergency management, with 83.33% This indicates that most of the research is geared towards applying theoretical AI concepts in real-world situations or analyzing existing case studies. In this context, Sahana EDEN is highlighted as a key solution in coordinating responses (Nussbaumer et al., 2023). A simulation game reflecting municipal council decisions during disasters offers insights into the complexity of disaster response (Adam & Lauradoux, 2022). Using drones for real-time accident monitoring is emphasized (El Barachi et al., 2020). AI's role in public health emergencies for efficient resource allocation (Wu & Shan, 2021). Also, VR is shown to be effective in Mass Casualty Incident (MCI) triage training Xia et al., 2022, while Sentinel-2 satellite imagery proves vital for environmental monitoring in natural disasters (Moumtzidou et al., 2020). The response to Hurricane Florence showcases practical technology applications (Ortiz et al., 2020). The STAMINA project illustrates AI's use in disaster management (Suciu et al.,

2021). The rapid damage assessment during Hurricane Dorian demonstrates the importance of swift evaluations post-disaster (Imran et al., 2020).

Additionally, AI's application for crisis management reflects its adaptability (Palomba et al., 2020). While lacking in detail, another study indicates AI's broad potential in disaster management (Karinshak, 2024). AI's practicality in real-world scenarios is also exemplified (Farasin et al., 2020), including its use in monitoring bridge health during natural disasters (Ramsey et al., 2023). Finally, AI's role in predicting disease spread is presented, underlining its importance in disaster planning and response (Landsberg et al., 2020). However, 16.67% of the references lack specific details, suggesting these discussions are more general or conceptual (Pandey et al., 2020; Sharma et al., 2019; Tsai et al., 2023).

EC6. Existing Regulatory Frameworks

The outcomes of our research reveal various dimensions in the study of using AI in emergency contexts within existing regulatory frameworks. Studies such as El Barachi et al., 2020, Wu and Shan, 2021, and De Nicola et al., 2019 emphasize adherence to data protection laws and the broader regulatory landscape encompassing AI in emergencies, accounting for three studies or 16.67% of the sample. The importance of aligning AI applications with regulations like the General Data Protection Regulation (GDPR) and German privacy laws is underscored in Farasin et al., 2020, Ramsey et al., 2023, and Landsberg et al., 2020, representing seven studies or 38.89%. In contrast, a focus on technical aspects, accounting for six studies or 33.33%, is seen in works that overlook in-depth regulatory discussions (Moumtzidou et al., 2020; Ortiz et al., 2020; Pandey et al., 2020; Sharma et al., 2019; Tsai et al., 2023; Xia et al., 2022). Nussbaumer et al., 2023 suggests that ethical imperatives should guide software development principles rather than directly addressing regulatory frameworks. Similarly, Adam and Lauradoux, 2022 focuses on the ethical and societal implications, prioritizing these over formal regulations. Jiang et al., 2021 addresses data privacy and user consent linked to regulatory compliance, while Karinshak, 2024 emphasizes ethical guidelines and advocates for a normative regulatory approach. Studies not specified in these categories account for two or 11.11% of the sample.

EC7. Ethical Considerations

A detailed analysis of ethical considerations in AI research is extensively covered by 18 studies, representing the entirety (100%) of this segment, including references This comprehensive coverage underscores the importance placed on ethical issues in AI research.

Privacy emerges as a crucial area of concern. A significant number of studies (12, accounting for 66.67%) delve into privacy issues, emphasizing the pervasive nature of privacy considerations in AI research highlight bias in data and algorithms, reflecting a critical awareness of potentially skewed interpretations and the need for unbiased computational methods (Adam & Lauradoux, 2022; Pandey et al., 2020; Wu & Shan, 2021).

Data handling and protection are also key areas of focus. Nine studies (50%) shed light on the importance of secure and ethical data management practices (De Nicola et al., 2019; Farasin et al., 2020; Jiang et al., 2021; Landsberg et al., 2020; Moumtzidou et al., 2020; Pandey et al., 2020; Suciu et al., 2021; Wu & Shan, 2021; Xia et al., 2022).

EC8. Model Used

A significant majority of the studies (14 out of 18, or 77.78%) specify the models they employ, including . These studies demonstrate a broad application of diverse computational techniques and models across different domains. In this sense, ML models and techniques are notably utilized, as indicated by studies (Ortiz et al., 2020; Pandey et al., 2020; Wu & Shan, 2021). Particularly, Wu and Shan, 2021 stands out for its implementation of neural network models and expert systems, highlighting the versatility of neural-based approaches and the reliance on structured knowledge.

Deep learning techniques, known for their depth and complexity, are employed in studies like (El Barachi et al., 2020; Jiang et al., 2021), reflecting their growing importance in extracting intricate patterns and insights from data. Karinshak, 2024 explores generative AI models for simulations and the broader application of generative AI, showcasing innovative uses of AI in creating dynamic and predictive environments. The role of AI in image processing is emphasized in Imran et al., 2020, enhancing the analysis and interpretation of visual data. Moreover, Palomba et al., 2020 presents the application of Random Forest models, exemplifying the use of ensemble learning methods for robust decision-making processes. On the other hand, a smaller proportion of studies (4 or 22.22%) do not specify the models used El Barachi et al., 2020; Nussbaumer et al., 2023; Tsai et al., 2023; Xia et al., 2022.

This diverse application of technologies, from ML to generative AI models, demonstrates the expansive reach and potential of AI and computational methods in advancing both research and practical applications.

Discussion

Strengths and Challenges

Integrating AI and ML in emergency management has highlighted several key strengths. A prominent one is the varied *AI Applications* in different contexts such as crisis scenario games and crowd sensing. Systems like Sahana EDEN are notable examples, significantly enhancing emergency response capabilities. In the realm of *Training*, technologies like Oculus Rift have revolutionized immersive training environments, while platforms like Google Dialogflow showcase the innovative use of AI in real-time analytics. *Data Analysis* has also seen advancements with tools like pyramid ConvLSTMs, demonstrating a commitment to developing sophisticated AI solutions. The *synergy between AI and Human Intelligence* is increasingly evident, with AI being combined with crowdsourced feedback for tasks such as image categorization. This synergy is further exemplified by applying models like Random Forest to satellite imagery. Lastly, a major strength lies in the *Transformation in Crisis Management*, where the use of deep learning on satellite data and the development of interactive visualizations for complex ML models enhance the interpretability of AI and lead to greater efficiency in emergency management.

However, the field also has notable challenges, representing key areas for future research and innovation. The unpredictability and variability of disasters present significant hurdles in developing robust and resilient AI-based response mechanisms. The user-friendliness and accessibility of serious gaming experiences, particularly for people with disabilities, is crucial for effective disaster education and remains an area needing improvement. Another challenge is ensuring the accuracy and timeliness of data for emergency responses, where AI support is essential yet still evolving. Careful attention to data collection and processing is necessary in virtual reality training integrated with AI. Another vital aspect is managing large datasets and ensuring their interoperability to improve disaster responses. Balancing AI precision with human judgment, especially in critical tasks like damage assessment, remains a crucial concern. Legal and data processing challenges, particularly in complex regulatory environments like Germany, also underline the need for harmonious legal frameworks and efficient data processing protocols. These challenges highlight the multifaceted nature of AI applications in emergency management and underscore the importance of ongoing research and development in addressing these issues effectively.

Highlighting the Potential for Future Research

Recent research in AI for emergency management underscores the importance of integrating ethics, enhancing communication, and ensuring interoperability. There is a growing need for further research into privacy-preserving AI techniques to bolster data analysis and establish secure data collection and sharing protocols with a strong emphasis on privacy. Advanced platforms in virtual environments are recommended to improve emergency response training and increase user engagement. The importance of integrating diverse data sources for better pattern recognition and chatbot refinement in disaster management is underscored, particularly in Indigenous communities, highlighting the need for cultural sensitivity in AI applications.

Continued research in AI ethics and regulation is crucial, focusing on enhancing model accuracy and efficiency. The expansion of VR and AI technologies for training and the development of broader AI tools for emergency management is advocated. Collaboration between humans and AI is vital, particularly in refining algorithms for various risk scenarios. The expansion of datasets and improvement of models are also essential, with an emphasis on ethical AI usage. Proposals for technological advancements include adopting cutting-edge data-cleaning technologies and improving visualization techniques to enhance data presentation. A major area of focus is increasing predictive accuracy in incident management and emergency vehicle routing, exploring AI's potential in areas such as predictive firefighting.

These recommendations collectively shed light on the future directions for integrating AI into emergency management, pointing towards a path of ethical and effective AI use in this critical field.

CONCLUSION

The integration of AI in the field of emergency management presents both groundbreaking opportunities and significant challenges. Various studies highlight the importance of incorporating ethics into information systems, improving communication and interoperability, and safeguarding privacy. It is recommended that serious games be used as educational tools for AI and its ethical implications and that further research into privacy-preserving AI techniques be conducted.

The findings suggest expanding AI applications in diverse cultural contexts and developing culturally sensitive AI solutions for disaster management. Expanding the use of VR and AI in training and developing AI tools for a wide range of emergency management applications is critical to enhancing the accuracy and efficiency of AI models.

For our future work, we plan to take a comprehensive approach to exploring the cultural, psychological, and logistical challenges of implementing AI in emergency settings. This will help us better understand how AI can effectively integrate into emergency management. In addition, our focus will be on identifying and investigating potential areas for further research and development. By doing so, we aim to extend the scope of our current study and improve its relevance and impact on AI applications in emergencies.

ACKNOWLEDGEMENTS

The work of Nu´n˜ez, A.G., Parra, O., and Alvear, O. was supported by *"Vicerrectorado de Investigacio´n e Innovacio´n (VIUC)"* at the University of Cuenca. The work of Nu´n˜ez, A.G., Quiroz-Palma, P., and Penade´s, M.C. was partially supported by grant MIGRATE (PID2019-105414RB-C31).

REFERENCES

Adam, C., & Lauradoux, C. (2022). A serious game for debating about the use of artificial intelligence during the covid-19 pandemic. *19th International Conference on Information Systems for Crisis Response and Management,{ISCRAM} 2022*.

Chen, B., Zhang, H., Sun, R., & Pan, J. (2023). Research on public satisfaction of government typhoon emergency management under artificial intelligence: An empirical analysis based on xuwen county. *AIMS Geosciences*, *9*(3), 466–491.

De Nicola, A., Melchiori, M., & Villani, M. L. (2019). Creative design of emergency management scenarios driven by semantics: An application to smart cities. *Information Systems*, *81*, 21–48.

El Barachi, M., Kamoun, F., Ferdaos, J., Makni, M., & Amri, I. (2020). An artificial intelligence based crowdsensing solution for on-demand accident scene monitoring. *Procedia Computer Science*, *170*, 303–310.

Farasin, A., Colomba, L., Palomba, G., Nini, G., Rossi, C., et al. (2020). Supervised burned areas delineation by means of sentinel-2 imagery and convolutional neural networks. *Proceedings of the 17th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2020), Virginia Tech, Blacksburg, VA, USA*, 24–27.

Freeman, S. (2020). Artificial intelligence for emergency management. *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*, *11413*, 393–400.

Imran, M., Alam, F., Qazi, U., Peterson, S., & Ofli, F. (2020). Rapid damage assessment using social media images by combining human and machine intelligence. *arXiv preprint arXiv:2004.06675*.

Jiang, R., Cai, Z., Wang, Z., Yang, C., Fan, Z., Chen, Q., Tsubouchi, K., Song, X., & Shibasaki, R. (2021). Deepcrowd: A deep model for large-scale citywide crowd density and flow prediction. *IEEE Transactions on Knowledge and Data Engineering*, *35*(1), 276–290.

Karinshak, E. (2024). Simulations in human-ai crisis management systems. *Journal of Contingencies and Crisis Management*, *32*(1), e12525.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. *Keele University*, *2.3*.

Landsberg, L., Ganske, D., Munschauer, C., & Mudimu, O. A. (2020). Using existing data to support operational emergency response in germany.

Mioch, T., Sterkenburg, R., Beuker, T., & Neerincx, M. A. (2021). Actionable situation awareness: Supporting team decisions in hazardous situations. *ISCRAM*, 62–70.

Moumtzidou, A., Bakratsas, M., Andreadis, S., Karakostas, A., Gialampoukidis, I., Vrochidis, S., & Kompatsiaris, I. (2020). Flood detection with sentinel-2 satellite images in crisis management systems. *ISCRAM 2020 Conference Proceedings-17th International Conference on Information Systems for Crisis Response and Management*, 1049–1059.

Nussbaumer, A., Pope, A., & Neville, K. (2023). A framework for applying ethics-by-design to decision support systems for emergency management. *Information Systems Journal*, *33*(1), 34–55.

Ortiz, B., Kahn, L., Bosch, M., Bogden, P., Pavon-Harr, V., Savas, O., & McCulloh, I. (2020). Improving community resiliency and emergency response with artificial intelligence. *arXiv preprint arXiv:2005.14212*.

Palomba, G., Farasin, A., Rossi, C., et al. (2020). Sentinel-1 flood delineation with supervised machine learning. *Proceedings of the 17th ISCRAM Conference*.

Pandey, R., Bannan, B., & Purohit, H. (2020). Citizenhelper-training: Ai-infused system for multimodal analytics to assist training exercise debriefs at emergency services. *ISCRAM 2020 conference proceedings–17th international conference on information systems for crisis response and management*.

Ramsey, A., Kale, A., Kassa, Y., Gandhi, R., & Ricks, B. (2023). Toward interactive visualizations for explaining machine learning models. *Proceedings of the Information Systems for Crisis Response and Management Conference, Omaha, NE, USA*, 28–31.

Sharma, S., Devreaux, P., Sree, S., Scribner, D., Grynovicki, J., & Grazaitis, P. (2019). Artificial intelligence agents for crowd simulation in an immersive environment for emergency response. *Electronic Imaging*, *2019*(2), 176–1.

Suciu, G., Pop, I. C., Birdici, A.-C., Ionescu, D., Vintila, A.-G., Matei, D., Ungureanu, A., & Modoaca, D. (2021). Stamina–assistive technology platform for pandemic prediction and crisis management. *2021 20th RoEduNet Conference: Networking in Education and Research (RoEduNet)*, 1–6.

Tsai, C.-H., Rayi, P., Kadire, S., Wang, Y.-F., Krafka, S., Zendejas, E., & Chen, Y.-C. (2023). Co-design disaster management chatbot with indigenous communities. *20th International Conference on Information Systems for Crisis Response and Management (ISCRAM), 2023*, 1.

Wu, Y., & Shan, S. (2021). Application of artificial intelligence to social governance capabilities under public health emergencies. *Mathematical problems in engineering*, *2021*, 1–10.

Xia, P., Ruan, J., Parry, D., Britnell, S., & Yu, J. (2022). Enhancing triage training for mass casualty incidents with virtual reality and artificial intelligence. *2nd Information Systems for Crisis Response and Management Asia Pacific Conference (ISCRAM 2022)*.

Yadav, A. (2024). Leveraging artificial intelligence for sustainable development and environmental resilience. In *Exploring ethical dimensions of environmental sustainability and use of ai* (pp. 140–165). IGI Global.

APPENDIX A

A list of papers selected for this Literature Review is included in Appendix A. Can be accessed in: Appendices-Resource.

Evacuation Social Support Among New Orleans Households: Implications for Evacuation Management Systems

Rob Grace
Texas Tech University
rob.grace@ttu.edu

Hyeong Suk Na
University of Missouri
hyeongsuk.na@missouri.edu

ABSTRACT

This study reports findings from a survey administered in 2023 to investigate New Orleans households' social support networks: family and friends who can provide them with accommodations and other assistance during evacuations caused by hurricanes and other hazards. Findings describe the demographics and evacuation destination and accommodation intentions of households with and without social support, as well as the factors these households consider when choosing an evacuation destination. Preliminary findings show that 86% of respondents report at least one relative or friend with whom they can stay during an evacuation. However, the range and size of respondents' social support networks vary and influences where they will go and stay during an evacuation. In contrast, 14% of respondents lack social support and these respondents are most likely to stay at a hotel or public shelter during an evacuation. Implications for the design of evacuation management systems are discussed, including systems that support the discovery of vulnerable households and coordination with community organizations that provide transportation and temporary housing during evacuations.

Keywords

Disaster management, evacuation, social capital, survey, vulnerability.

INTRODUCTION

This work-in-progress study examines the social support networks of Greater New Orleans households: family and friends who can provide them with accommodations and other assistance during an evacuation. While prior survey research examines where households travel during an evacuation, and whether they stay at the homes of family and friends, hotels, or public shelters (Lindell et al., 2019), these studies do not examine the extent to which households can obtain accommodations with friends and relatives, and the demographic characteristics and evacuation intentions of households with and without social support. Lacking insight into households' social support networks presents two problems for research attempting to understand large-scale evacuation behavior and design information systems that can support household and community evacuation planning.

First, survey studies show that evacuees overwhelmingly rely on family and friends for temporary accommodations, but seldom examine households' social support networks to identify vulnerable groups who, lacking assistance from family and friends, may require assistance from community organizations during an evacuation. Surveys of intended and actual accommodations show that the majority of evacuees stay at the homes of family and friends, fewer stay in hotels, and far fewer stay in public shelters (Lindell et al., 2011; Mesa-Arango et al., 2013; Mitchell et al., 2017; Wong et al., 2018; Wu et al., 2012; Zeigler et al., 1981). Although emerging studies highlight unequal access to social support among households (Grace & Na, 2024; Na & Grace, 2022), empirical studies of households' social support networks are needed to characterize vulnerable groups that likely require transportation and accommodation assistance from public authorities and community groups during an evacuation.

Second, survey studies show that evacuees travel to diverse in-county, in-state, and out-of-state destinations (Morrow & Gladwin, 2005; Smith & McCarty, 2009; Wong et al., 2018; Wu et al., 2012), but do not adequately

WiP Paper – Open Track
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

explain why households travel to one destination over another. Beyond hazard proximity and related risk perceptions, such as residence in coastal counties threatened by hurricanes (Huang et al., 2016), studies observe weak or inconsistent correlations between evacuation destination and demographic variables such as age, disability, home ownership, income, race, and social vulnerability (Adeola & Picou, 2017; Asfaw et al., 2019; Baker, 1991; Bolin & Kurtz, 2018; Wu et al., 2012; Yabe & Ukkusuri, 2020). Understanding where households can travel to stay with family and friends can provide insight into households' likely evacuation destinations and, in turn, predict demand for accommodations such as hotels and public shelters.

These gaps in our knowledge of households' social support networks limit, in turn, our ability to design decision-support, simulation, and geographic information systems that can help emergency managers plan community evacuations by, for example, predicting the time of departure, routes of travel, destinations, and types of accommodations evacuees will require after leaving an evacuation zone (Lindell & Prater, 2007a; Yang et al., 2019). Simulating evacuation behavior requires empirically validated models of household evacuation destination and accommodation decision-making (Lindell et al., 2019; Murray-Tuite & Wolshon, 2013). These models and simulations support emergency management planning that involves coordinating public transportation and locating public shelters. Furthermore, knowledge of social support network availability and accessibility can inform the design of systems that can help identify and connect vulnerable households with transportation services and temporary accommodations.

To address these gaps in the literature, we report findings from an intended response survey administered to households in the Greater New Orleans area (n=305). Survey respondents were asked where they could go to stay with family or friends during an evacuation, where they would go and stay in the event of a future evacuation, and what factors they consider when choosing an evacuation destination and accommodations. We answer three research questions in this work-in-progress paper:

1. *To what extent can households stay at the homes of family or friends during an evacuation?* We identify households with and without social support networks, i.e., family and friends who can provide them with temporary accommodations during an evacuation.
2. *What are households' intended evacuation destinations and accommodations?* We compare where households with and without social support intend to travel and stay during an evacuation.
3. *What factors do households consider when choosing an evacuation destination?* We compare the relative importance of decision-making factors such as cost, transportation, and destination familiarity to households with and without social support.

The following sections describe, first, prior research on the relationship between social support and evacuation intentions and behaviors. Second, we describe the methods of survey design and data analysis. Lastly, we report our results in answer to the three research questions above and discuss implications for technologies that can support evacuation planning.

BACKGROUND

Prior research informs our knowledge of the availability of social support and the relationship between social support and evacuation intentions and behaviors. This research includes studies that examine households' (i) social support and related concepts like social capital, (ii) accommodations, and (iii) destinations observed during actual and planned evacuations from areas threatened by hurricanes, wildfires, nuclear accidents, and other disasters.

Definition of Social Support

Social support refers to informational, material (e.g., accommodations), and emotional assistance that households receive or expect to receive from their friends and relatives, i.e., social support network, during an evacuation (Asfaw et al., 2019). Studies of evacuation have long realized the importance of social support networks during evacuations. Quarantelli (1985), for instance, discusses "social support systems" as the family and friends who provide evacuees with temporary shelter (short-term accommodations) and housing (long-term accommodations), observing that the characteristics of households' social support systems (what we refer to as social support networks) make them more or less resilient when forced to remain away from home for varying lengths of time.

The concepts of social support and social support networks respectively distinguish the type and source of assistance that households receive or expect during an evacuation. By doing so, they unpack individual-level measures of social capital (Meyer, 2018), which can describe evacuees' "access to and use of resources embedded in social networks" (Lin, 2001, p. 5). In this study, we refer to social support networks and social support rather than social capital to understand the geographic distribution of households' family and friends and the types of

resources these social network ties can locally provide to households, respectively. While households can receive information and some material assistance (e.g., money) from remote family and friends, in this study we focus on the provision of local assistance, namely accommodations.

Social Support and Evacuation Accommodations

Accommodations are arguably the most important and common type of material assistance that evacuees receive from their social support networks. Prior survey studies examine where respondents traveled during a past evacuation (i.e., actual response) or where they intend to travel for a future evacuation (i.e., intended response). Actual response surveys typically find 50-90% of evacuees staying with family and friends, fewer staying in hotels or motels, and the fewest staying at public shelters (Table 1). These studies also show that evacuees tend to stay with relatives versus friends (Siebeneck & Cova, 2012; Wong et al., 2018). Wong et al. (2018), for instance, surveyed 368 evacuees from Hurricane Irma to find that 44% of respondents stayed with relatives while only 16% stayed with friends. Additionally, these studies note that some respondents stayed at "other" locations such as peer-to-peer rentals (e.g., Airbnb, etc.), churches, and workplaces (Mesa-Arango et al., 2013; Wong et al., 2018).

Table 1. Actual and intended (in gray) evacuation accommodations of U.S. survey respondents.

Disasters	Family/Friends	Hotel	Shelter	Other	Unknown
Three Mile Island nuclear accident in PA (Brunn et al., 1979)[1]	81%	17%	0%	4%	
Hypothetical nuclear accident in PA (Brunn et al., 1979)[1]	70%	15%	9%	5%	
Hurricane Bonnie in NC (Whitehead et al., 2000)	70%	16%	5%	9%	
Future hurricane in NC (Whitehead et al., 2000)	60%	24%	12%	4%	
Hurricane Katrina in LA (Wu et al., 2012)	54%	31%	3%	13%	
Hurricane Katrina in LA (Mitchell et al., 2017)[1]	54%	30%	6%	9%	
Future hurricane in LA (Mitchell et al., 2017)	51%	33%	5%	8%	3%
Future hurricane in LA (Morrow & Gladwin, 2006)	57%	27%	5%	3%	8%
Hurricane Harvey in TX (Bierling et al., 2020)	58%	33%	1%	8%	

[1]*Values reported in the original study do not equal 100%.*

Intended response surveys ask respondents about their planned evacuation accommodations and destinations for future evacuation scenarios, such as a "major hurricane like Katrina" (Morrow & Gladwin, 2006). Table 1 also describes the findings of intended response surveys administered to U.S. households threatened by hurricanes and, in the case of Three Mile Island, the partial meltdown of a nuclear reactor. These and other studies suggest that respondents' accommodation intentions correspond with actual accommodation decisions observed during prior or future evacuations (Kang et al., 2007). For instance, in a survey of Southeastern Louisiana residents, Mitchell et al. (2017) observe that respondents' accommodation intentions for a future hurricane generally match the actual accommodations they obtained during Hurricanes Katrina and Gustav.

Studies also investigate the relationship between evacuees' actual and intended accommodations and demographic and evacuation factors. Some studies observe that low-income households with young children, elderly persons, or persons with disabilities/special needs are more likely to evacuate to homes of family or friends or stay at a public shelter than stay at a hotel (Whitehead et al., 2000; Wong et al., 2018). Studies also observe that whites, large families, and high-income households are more likely to stay in hotels and less likely to stay in shelters (Lindell et al., 2011; Wu et al., 2012). Lastly, other studies observe a relationship between households' home city/state and choice of accommodations. For example, Wu et al. (2012) find that Louisiana households were more likely to stay in hotels than Texas households during Hurricane Katrina.

Evacuation Destinations

Studies show that most U.S. residents of the Gulf and Atlantic Coasts evacuate to in-state and out-of-state

destinations in response to hurricanes, but some evacuate to destinations within their home county or parish (Lindell & Prater, 2007b). Table 2 reports the actual destinations of surveyed evacuees, showing that respondents reported a wide range of in-county (1-49%), in-state (30-81%), and out-of-state (0-56%) destinations. These measures vary significantly depending on the geographic area surveyed. Dash and Morrow (2000), for example, surveyed residents of the Florida Keys and observed that none evacuated the 400+ miles required to leave the state of Florida. In contrast, many New Orleans residents surveyed by Mitchell et al. (2017) and Wu et al. (2012) traveled to relatively close out-of-state destinations in Texas and Mississippi.

Table 2. Actual and intended (in grey) evacuation destinations of U.S. survey respondents.

Evacuation	Evacuation destinations		
	In county	**In state**	**Out of state**
Hurricane Georges in FL (Dash & Morrow, 2000)	19%	81%	0%
Hurricane Floyd in SC (Dow & Cutter, 2002)	9%	32%	56%
Hurricane Ivan (Morrow & Gladwin, 2005)	32%	37%	31%
Multiple hurricanes in FL (Smith & McCarty, 2009)	49%	30%	21%
Hurricane Katrina (Wu et al., 2012)	1%	80%	19%
Hurricane Rita (Wu et al., 2012)	1%	47%	52%
Hurricane Katrina/Gustav/Isaac (Mitchell et al., 2017)	14%	34%	52%
Hurricane Irma (Wong et al., 2018)	17%	34%	49%
Future hurricane in LA (Mitchell et al., 2018)[1]	-	27%	73%

[1]*Did not report in-county evacuation destinations.*

Lastly, research explores relationships between households' evacuation destinations and demographic and non-demographic factors such as social support. In general, demographic factors such as age, disability, home ownership, income, race, and social vulnerability fail to predict evacuation destinations and distances (Adeola and Picou, 2017; Asfaw et al., 2019; Bolin and Kurtz, 2018; Howard et al., 2017; Yabe and Ukkusuri, 2020; Wu et al., 2012). However, emerging studies that examine social support from family and friends observe that evacuees plan to travel shorter distances when evacuating to destinations where they can receive assistance from family and friends but travel farther when planning to evacuate to destinations lacking social support (Na & Grace, 2022).

METHODS

Although prior evacuation research examines the accommodations and destinations of households fleeing natural hazards like hurricanes, a gap exists in our understanding of social support available to households and how the availability of social support influences where households choose to go and stay during an evacuation. To address this gap, we surveyed Greater New Orleans households about their social support networks and evacuation intentions. We describe the design of the survey and methods of data analysis below. This survey research was approved by a university institutional review board (IRB2023-735).

The survey was administered to an online panel (n=305) composed of Southeast Louisianna residents mostly located in New Orleans Parish, but including respondents from surrounding areas, including St. Charles, Jefferson, and St. Tammany Parishes. Respondents completed a 34-question survey that included demographic, evacuation, and social support-related questions. These questions were based on prior survey research on evacuation behavior (Bierling et al., 2020; Mitchell et al., 2017) and FEMA's National Household Survey (2023).

Our analysis of the survey data addresses the following research questions: (RQ1) To what extent can households stay at the homes of family or friends during an evacuation? (RQ2) What are households' intended evacuation accommodations and destinations? (RQ3) What factors do households consider when choosing an evacuation destination? We cross-tabulated the survey data to understand the number and demographics (RQ1), accommodation and destination intentions (RQ2), and destination decision-making factors (RQ3) of respondents with and without social support.

RESULTS

In this section, we report survey findings to compare the demographics, evacuation intentions, and destination decision-making factors of respondents with and without social support networks that vary in range and size.

Availability of Social Support

Of 305 survey respondents, 263 (86%) identified at least one relative or friend who could provide them with accommodations during an evacuation, while 42 (14%) did not identify a source of social support (Table 1). These two groups represent households for whom social support is available and unavailable.

The demographics of respondents with and without social support are roughly similar (Table 4). Respondents with social support are more likely to be female, married, and in early and middle adulthood (25-65), whereas respondents without social support are more likely to be male, unmarried, and elderly (65+). Additionally, there are more respondents with social support who earn less than $55k or more than $100k annually than respondents without social support who are more likely to be middle-income earners. Respondents with social support tend to have more evacuation experience than respondents without social support. Importantly, however, the two groups are roughly similar in terms of race, educational attainment, employment status, and household size.

Table 4. Demographics of respondents with and without social support.

Characteristic		All (n=305)	Social Support	
			Available (n=263)	Unavailable (n=42)
Age	18-24	39 (13%)	33 (13%)	6 (14%)
	25-65	204 (67%)	183 (70%)	21 (50%)
	65+	62 (20%)	47 (18%)	15 (36%)
Gender	Female	199 (65%)	175 (67%)	24 (57%)
	Male	105 (34%)	87 (33%)	18 (43%)
Race	Black	92 (30%)	78 (30%)	14 (33%)
	White	183 (60%)	158 (60%)	25 (60%)
	Other	30 (10%)	27 (10%)	3 (7%)
Relationship status	Married	99 (32%)	90 (34%)	9 (21%)
	Single	92 (30%)	80 (30%)	12 (29%)
	In a relationship	50 (16%)	41 (16%)	9 (21%)
	Other (e.g., widowed)	64 (21%)	52 (20%)	12 (29%)
Educational attainment	High school or less	126 (41%)	109 (41%)	17 (40%)
	Associate degree or vocational training	79 (26%)	66 (25%)	13 (31%)
	Bachelor's degree or higher	100 (33%)	88 (33%)	12 (29%)
Employment status	Unemployed	33 (11%)	30 (11%)	3 (7%)
	Employed (full or part-time)	169 (55%)	146 (56%)	23 (55%)
	Student	19 (6%)	16 (6%)	3 (7%)
	Retired	84 (28%)	71 (27%)	13 (31%)
Income	$0-49k	148 (49%)	135 (51%)	13 (31%)
	$50-99K	102 (33%)	79 (30%)	23 (55%)
	$100k+	55 (18%)	49 (19%)	6 (14%)
Prior evacuation experience	No	56 (18%)	45 (17%)	11 (26%)
	Yes	249 (82%)	218 (83%)	31 (74%)
Persons in household	1 person	90 (30%)	76 (29%)	14 (33%)
	2-4 persons	183 (60%)	157 (60%)	26 (62%)
	5+ persons	32 (10%)	30 (11%)	2 (5%)

Accommodation Intentions

Asked where they would stay in the event of a future one-week evacuation, respondents with and without social support indicated they would stay in different types of accommodations. Table 5 shows the intended accommodations of respondents according to the range of their social support networks, i.e., the in-county, in-

state, or out-of-state location of the most distant family or friend who can provide them with a place to stay during an evacuation. Respondents with social support are most likely to stay family or friends (58%), however, respondents with out-of-state social support (61%) are more likely to stay with family or friends than respondents with in-county (42%) or in-state (49%) social support. However, respondents with only local, in-county social support are more likely to stay at hotels (42%) than respondents with more expansive social support networks. In contrast, respondents without social support are most likely to stay at hotels (60%) or public shelters (21%).

Table 5. Intended accommodations of respondents by range of social support network.

	Social Support		Range of Social Support Network		
Accommodation	**Unavailable**	**Available**	**In County[1]**	**In State**	**Out of State**
House of family/friend	0 (0.0%)	152 (57.8%)	8 (42.1%)	19 (48.7%)	125 (61.0%)
Hotel or motel	25 (59.5%)	75 (28.5%)	8 (42.1%)	10 (25.6%)	57 (27.8%)
Other (RV, 2nd home, etc.)	5 (11.9%)	15 (5.7%)	2 (10.5%)	6 (15.4%)	7 (3.4%)
Peer-to-peer rental	3 (7.1%)	16 (6.1%)	0 (0.0%)	2 (5.1%)	14 (6.8%)
Public shelter	9 (21.4%)	5 (1.9%)	1 (5.3%)	2 (5.1%)	2 (1.0%)
Overall	42 (100.0%)	263 (100%)	19 (100.0%)	39 (100.0%)	205 (100.0%)

[1]For comparison with prior research that reports in-county destinations, we refer to counties rather than parishes.

Table 5 also reveals that some respondents who have social support choose not to use it. Among 263 respondents with at least one family or friend with whom they can stay during an evacuation, 111 (42%) intend to stay in a hotel or other accommodations. Not shown in Table 5, 32 (12%) respondents plan to stay in a hotel or other accommodation at a destination where they could stay at the home of a friend or relative. These "locally supported" households have not been identified before in prior survey studies.

Table 6 shows the intended accommodations of respondents according to the size of their social support networks, i.e., the number of family and/or friends who can provide them with a place to stay during an evacuation. Respondents with larger social support networks are more likely to stay with family and friends: over 50% of respondents with social support networks of 1-2 persons intend to stay with these family and friends, while 100% of respondents with three persons in their social support networks intend to stay in these accommodations. In contrast, as reported in Table 5, most respondents without social support intend to stay in hotels or public shelters.

Table 6. Intended accommodations of respondents by size of social support network.

	Size of Social Support Network		
Accommodation	**1 Person**	**2 Persons**	**3 Persons**
House of family/friend	34 (50.7%)	85 (52.1%)	33 (100%)
Hotel or motel	22 (32.8%)	53 (32.5%)	0 (0%)
Other (RV, second home, etc.)	4 (6.0%)	11 (6.7%)	0 (0%)
Peer-to-peer rental	4 (6.0%)	12 (7.4%)	0 (0%)
Public shelter	3 (4.5%)	2 (1.2%)	0 (0%)
Total	67 (100%)	163 (100%)	33 (100%)

Destination Intentions

Table 7 shows the intended evacuation destinations of respondents according to the range of their social support networks. Among 305 respondents, 205 (67%) can travel out of state to stay with a family/friend, 39 (13%) can travel somewhere within the state of Louisiana, 19 (6%) can only access social support within their home parish, and 42 (14%) have nowhere to go for help from family/friends.

Significantly, and related to the preference of staying at the homes of family/friends, where respondents plan to evacuate relates to the range of their social support networks. Households with only in-county social support tend to remain within the county/parish (58%), households with in-state social support tend to evacuate within the state (49%), and households with out-of-state social support tend to evacuate to destinations in the neighboring states of Alabama, Mississippi, and Texas (85%). Importantly, unlike respondents with out-of-state support, respondents with no support (24%), in-county support (16%), and in-state support (13%) are more likely to not know where they would evacuate.

Table 7. Intended evacuation destinations by range of social support network.

	Social Support		Range of Social Support Network		
Evacuation Destination	**Unavailable**	**Available**	**In County**	**In State**	**Out of State**
In County	7 (16.7%)	14 (5.3%)	11 (57.9%)	1 (2.6%)	2 (1.0%)
In State	2 (4.8%)	40 (15.2%)	1 (5.3%)	19 (48.7%%)	20 (9.8%)
Out of state	23 (54.8%)	193 (73.4%)	4 (21.1%)	14 (35.9%%)	175 (85.4%)
Unknown	10 (23.8%)	16 (6.1%)	3 (15.8%)	5 (12.8%%)	8 (3.9%)
Total	42 (100%)	263 (100%)	19 (100%)	39 (100%)	205 (100%)

Table 7 also shows that respondents with local support often evacuate outside of their social support networks by planning to travel to out-of-state destinations where they intend to stay at a hotel or other accommodation. Among 39 respondents with in-state social support, 14 (36%) intend to evacuate outside Louisiana where they lack family/friends who can provide them with accommodations. Overall, most respondents plan to evacuate to an out-of-state destination (Figure 1).

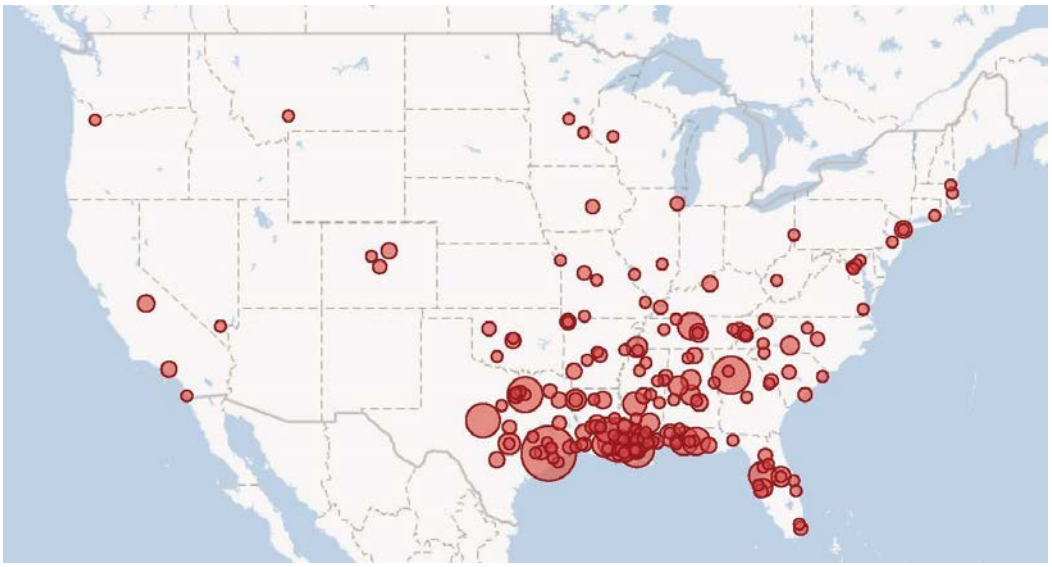


Figure 1. Intended destinations of survey respondents.

Table 8 shows the intended evacuation destinations of respondents according to the size of their social support networks. While out-of-state destinations are the most common among respondents with 1-3 family and/or friends who can provide them with a place to stay during an evacuation, respondents with more family/friends are more likely to evacuate out of state than respondents with smaller social support networks. Approximately 88% of respondents with 3 family/friends intend to travel out of state compared to 69% of respondents with one family/friend and 55% of respondents without social support.

Table 8. Intended evacuation destinations by size of social support network.

	Size of Social Support Network		
Evacuation Destination	**1 FF**	**2 FF**	**3 FF**
In County	6 (9.0%)	7 (4.3%)	1 (3.0%)
In State	11 (16.4%%)	27 (16.6%)	2 (6.1%%)
Out of State	46 (68.7%%)	118 (72.4%)	29 (87.9%%)
Unknown	4 (6.0%%)	11 (6.7%)	1 (3.0%%)
Total	67 (100%)	163 (100%)	33 (100%)

Decision-making Factors

Table 8 shows the factors that respondents consider when choosing an evacuation destination. Among the nine factors surveyed, respondents most often identified accommodation availability, destination distance, and financial

cost as important when deciding where to go during a future evacuation. However, respondents with social support (68%) are more likely to consider the availability of accommodations when choosing an evacuation destination than respondents without social support (43%). Furthermore, respondents with social support are more likely to consider the recommendations of family and friends (19%) than unsupported respondents (5%), as well as their familiarity with the destination (32%) compared to unsupported respondents (19%). In contrast, respondents without social support are more likely to consider cost (55%) and transportation (48%) to destinations than respondents with social support (49% and 38% respectively).

Table 8. Factors considered when choosing an evacuation destination.

	Social Support		Range of Social Support		
Factor	**Unavailable**	**Available**	**Local**	**State**	**Out of State**
Accommodations	18 (43%)	179 (68%)	10 (53%)	22 (56%)	119 (58%)
Distance	21 (50%)	151 (57%)	11 (58%)	24 (62%)	144 (70%)
Cost	23 (55%)	130 (49%)	12 (63%)	10 (26%)	77 (38%)
Transportation	20 (48%)	99 (38%)	7 (37%)	18 (46%)	71 (35%)
Traffic congestion	14 (33%)	96 (37%)	5 (26%)	16 (41%)	109 (53%)
Familiarity	8 (19%)	83 (32%)	5 (26%)	6 (15%)	38 (19%)
Family/friends advice	2 (5%)	49 (19%)	3 (16%)	5 (13%)	33 (16%)
Officials/media advice	4 (10%)	41 (16%)	3 (16%)	12 (31%)	67 (33%)
Other	4 (10%)	13 (5%)	1 (5%)	0 (0%)	12 (6%)

DISCUSSION

The preliminary survey results provide insight into the availability of social support among Greater New Orleans households and the relationship between social support, namely a place to stay with family and friends, and households' evacuation accommodation and destination intentions. The first of its kind, these results show that:

- Among survey respondents from Greater New Orleans (n=305), 263 (86%) report at least one friend or relative at whose home they can stay during an evacuation, however, 42 (14%) report no sources of social support.
- The ranges of social support networks vary among respondents (n=263): 19 (7%) report only friend/relative(s) inside their home parish, 39 (15%) report someone elsewhere in the state of Louisiana, and 205 (78%) report someone outside the state.
- The sizes of social support networks vary among respondents (n=263): 67 (22%) report one friend/relative with whom they can stay during an evacuation, 163 (53%) report two persons, and 33 (11%) report three persons.
- Respondents with social support are most likely to stay at the home of a friend or relative (58%), while respondents without social support are most likely to stay at a hotel (60%) or public shelter (21%). In contrast, 29% of respondents with social support intend to stay at a hotel and 2% intend to stay at a public shelter.
- Although they might stay with a friend or relative, 42% of respondents with social support intend to stay at a hotel, peer-to-peer rental, RV, public shelter, or other accommodation.
- Most respondents with and without social support intend to evacuate to out-of-state destinations, but respondents with social support are most likely to evacuate to in-parish, in-state, and out-of-state destinations where they can stay at the homes of family and friends.
- Respondents with social support are more likely to consider the availability of accommodations and recommendations of friends/relatives when choosing an evacuation destination than respondents without social support.

These results, by describing the availability, range, and size of households' social support networks, provide context to prior survey research that examines actual and intended evacuations destinations and accommodations (Adjei et al., 2022; Bierling et al., 2020; Huang et al., 2016; Lindell et al., 2011; Mitchell et al., 2017; Smith & McCarty, 2009; Wong et al., 2018; Wu et al., 2012). Importantly, the findings suggest that differences in social support networks may be better indicators of household vulnerability than demographic and other risk factors observed in previous studies (Adeola & Picou, 2017; Asfaw et al., 2019; Baker, 1991; Bolin & Kurtz, 2018; Wu

et al., 2012; Yabe & Ukkusuri, 2020).

By describing the friends/relatives who can provide households with accommodations during an evacuation, the preliminary results suggest that the range/geographic distribution and size of households' social support networks influences where they go and stay during an evacuation. Consequently, by improving our understanding of households' social support networks, we can better explain and predict where people go during an evacuation, their logistics requirements, and identify vulnerable groups who lack social support or the means to access social support during a large-scale evacuation.

Limitations of the present study include the non-representative sample of Greater New Orleans residents, biases introduced through the use of an online panel to administer the survey, and possible differences between respondents intended and actual response to an evacuation (Kang et al., 2007; Lindell et al., 2019). Future studies can address these limitations through different survey designs and recruitment procedures, as well as complementary methods, such as interviews, that can help explain the patterns of intentional and actual behavior observed through survey research. As described in the next section, future research can also explore the implications of these findings for the design of evacuation management systems.

Implications for Evacuation Management Systems

Information systems can support the management of large-scale evacuations at community and household levels. This study and future research can inform the design of such systems by describing and predicting household vulnerability and exposure to hurricane and evacuation-related hazards based on the availability and accessibility of households' social support. In addition to often-studied technologies, such as decision-support systems (DSS) for emergency managers (Davidson et al., 2020), this study suggests the need for systems that help households plan to access social support for likely evacuation scenarios, as well as seek assistance from government and non-government agencies (NGOs) when social support is unavailable or inaccessible during an evacuation.

At the community level, prior research focuses on evacuation and transportation modeling and simulation, and the design of DSS and intelligent transportation systems (ITS) for emergency managers and other decision-makers responsible for planning and managing evacuation and return-entry processes (Davidson et al., 2020; Iliopoulou et al., 2020; Murray-Tuite & Wolshon, 2013). As Lindell and Prater (2007a) explain, emergency managers in New Orleans and other hurricane-threatened areas depend on DSS for insight into "the behavior of the hurricane that is relevant to evacuation and the behavior of evacuees that is relevant to the hurricane" (p. 627). By providing insight into households' social support networks and evacuation destination and accommodation intentions, this work-in-progress paper and future research can inform models of evacuation behavior that, in turn, can inform the design of DSS that can guide emergency managers in evacuation planning, including the allocation and location of public shelters (Lindell & Prater, 2007a; Yang et al., 2019).

However, our findings also suggest opportunities for systems that support the discovery of vulnerable households that lack social support to connect these households with government agencies and/or NGOs poised to help during an evacuation. For instance, our preliminary results suggest that Greater New Orleans residents who are male, unmarried, and elderly are more likely to lack social support than other demographic groups. Designs might include social media listening and public information campaigns that target users who likely lack social support to direct them to web-based resources for obtaining transportation and accommodations ahead of a possible evacuation. In New Orleans, for example, city-assisted evacuation coordinates transportation out of the city using a system of pre-established pick-up areas and designated evacuation destinations with public shelters. Importantly, residents need to reach these pick-up points with the proper provisions and at the proper times. Future research can partner with community partners to co-design systems that identify and initiate communications with vulnerable households to effectively coordinate the use of such public transportation systems.

At the household level, insights from this study suggest the need for systems that support evacuation planning for different hurricane scenarios, particularly those in which a household may not be able to access accommodations from family or friends. This study finds that Greater New Orleans households have social support networks that vary in range and size, which suggests that the exposure of households' social support networks to hurricane hazards and evacuation orders will vary depending on the location of households' family and friends. Consequently, households may need to stay at a hotel or public shelter when social support is unavailable, as observed among the 14% of respondents who could not identify a friend or relative with whom they could stay during an evacuation, or inaccessible, such as when the path of a storm or road conditions restrict travel to available homes of friends and relatives.

As few households proactively plan for an evacuation (Vásquez et al., 2016), systems might support households in contingency planning for scenarios when they may need to stay in a hotel or other accommodations when social

support is available but inaccessible. Similar to DSSs emergency managers use to predict evacuation zones based on likely storm paths and intensities (Lindell et al., 2019), a system might help households plan where to go and stay based on the availability and accessibility of social support during likely hurricane scenarios. As DSS research has not focused on household use contexts, future work is required to understand households' evacuation planning needs and co-design and evaluate systems that can support destination, accommodations, and logistics planning ahead of future disasters.

CONCLUSION

This study reports preliminary findings from a survey administered in 2023 to investigate New Orleans households' social support networks: friends and relatives who can provide them with accommodations and other assistance during evacuations caused by hurricanes and other hazards. While most respondents report at least one relative or friend with whom they can stay during an evacuation, some respondents lack social support, and these respondents are most likely to stay at a hotel or public shelter during an evacuation. These findings and future research can improve understanding of the relationship between social support and evacuation behavior and inform the development of systems that can support community and household evacuation planning.

REFERENCES

Adeola, F. O., & Picou, J. S. (2017). Hurricane Katrina-linked environmental injustice: Race, class, and place differentials in attitudes. *Disasters*, *41*(2), 228–257. https://doi.org/10.1111/disa.12204

Adjei, E., Murray-Tuite, P., Ge, Y., Ukkusuri, S., & Lee, S. (2022). Estimating hurricane evacuation destination and accommodation type selection with perceived certainty variables. *Transportation Research Part D: Transport and Environment*, *105*, 103235. https://doi.org/10.1016/j.trd.2022.103235

Asfaw, H. W., McGee, T., & Christianson, A. C. (2019). The role of social support and place attachment during hazard evacuation: the case of Sandy Lake First Nation, Canada. *Environmental Hazards*, *18*(4), 361–381. https://doi.org/10.1080/17477891.2019.1608147

Baker, E. J. (1991). Hurricane evacuation behavior. *International Journal of Mass Emergencies and Disasters*, *9*(2), 287–310.

Bierling, D. H., Lindell, M. K., Peacock, W. G., Abuabara, A., Moore, R. A., Wunneburger, D. F., Mullins, J. A., & Borchardt, D. W. (2020). *Coastal Bend Hurricane Evacuation Study: Hurricane Harvey Evacuation Behavior Survey Outcomes and Findings*. https://hdl.handle.net/1969.1/188203

Bolin, B., & Kurtz, L. C. (2018). *Race, Class, Ethnicity, and Disaster Vulnerability* (pp. 181–203). https://doi.org/10.1007/978-3-319-63254-4_10

Dash, N., & Morrow, B. H. (2000). Return delays and evacuation order compliance: the case of Hurricane Georges and the Florida Keys. *Environmental Hazards*, *2*(3), 119–128. https://doi.org/10.3763/ehaz.2000.0217

Davidson, R. A., Nozick, L. K., Wachtendorf, T., Blanton, B., Colle, B., Kolar, R. L., DeYoung, S., Dresback, K. M., Yi, W., Yang, K., & Leonardo, N. (2020). An Integrated Scenario Ensemble-Based Framework for Hurricane Evacuation Modeling: Part 1—Decision Support System. *Risk Analysis*, *40*(1), 97–116. https://doi.org/10.1111/risa.12990

Dow, K., & Cutter, S. L. (2002). Emerging Hurricane Evacuation Issues: Hurricane Floyd and South Carolina. *Natural Hazards Review*, *3*(1), 12–18. https://doi.org/10.1061/(ASCE)1527-6988(2002)3:1(12)

FEMA. (2023). *FEMA National Household Survey*. https://www.fema.gov/about/openfema/data-sets/national-household-survey

Grace, R., & Na, H. S. (2024). Relocation and Social Support during Large-Scale Evacuations. *Natural Hazards Review*, *25*(2), 1–17. https://doi.org/10.1061/NHREFO.NHENG-1874

Huang, S. K., Lindell, M. K., & Prater, C. S. (2016). Who Leaves and Who Stays? A Review and Statistical Meta-Analysis of Hurricane Evacuation Studies: *Environment and Behavior*, *48*(8), 991–1021. https://doi.org/10.1177/0013916515578485

Iliopoulou, C., Konstantinidou, M. A., Kepaptsoglou, K. L., & Stathopoulos, A. (2020). ITS Technologies for Decision Making during Evacuation Operations: A Review. *Journal of Transportation Engineering, Part A: Systems*, *146*(4). https://doi.org/10.1061/JTEPBS.0000329

Kang, J. E., Lindell, M. K., & Prater, C. S. (2007). Hurricane Evacuation Expectations and Actual Behavior in Hurricane Lili. *Journal of Applied Social Psychology*, *37*(4), 887–903. https://doi.org/10.1111/j.1559-1816.2007.00191.x

Lin, N. (2001). Building a network theory of social capital. In N. Lin, K. Cook, & R. S. Burt (Eds.), *Social capital: Theory and research* (pp. 3–28). Transaction Publishers.

Lindell, M. K., Kang, J. E., & Prater, C. S. (2011). The logistics of household hurricane evacuation. *Natural Hazards*, *58*(3), 1093–1109. https://doi.org/10.1007/s11069-011-9715-x

Lindell, M. K., Murray-Tuite, P., Wolshon, B., & Baker, E. J. (2019). *Large-scale evacuation: The analysis, modeling, and management of emergency relocation from hazardous areas*. Routledge.

Lindell, M. K., & Prater, C. S. (2007a). A hurricane evacuation management decision support system (EMDSS). *Natural Hazards*, *40*(3), 627–634. https://doi.org/10.1007/s11069-006-9013-1

Lindell, M. K., & Prater, C. S. (2007b). Critical Behavioral Assumptions in Evacuation Time Estimate Analysis for Private Vehicles: Examples from Hurricane Research and Planning. *Journal of Urban Planning and Development*, *133*(1), 18–29. https://doi.org/10.1061/(ASCE)0733-9488(2007)133:1(18)

Mesa-Arango, R., Hasan, S., Ukkusuri, S. V., & Murray-Tuite, P. (2013). Household-Level Model for Hurricane

Evacuation Destination Type Choice Using Hurricane Ivan Data. *Natural Hazards Review*, *14*(1), 11–20. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000083

Meyer, M. A. (2018). *Social Capital in Disaster Research* (pp. 263–286). https://doi.org/10.1007/978-3-319-63254-4_14

Mitchell, B., Meyer, M., Henderson, M., Barnes, S., Morrow, B., Gladwin, H., Smith, J., Breen, K., & Iles, R. (2017). *Behavioral analysis for Southeast Louisiana hurricane events*.

Morrow, B. H., & Gladwin, H. (2005). *Hurricane Ivan behavioral analysis*.

Morrow, B. H., & Gladwin, H. (2006). *Coastal Louisiana Post-Katrina Evacuation Planning Survey*.

Murray-Tuite, P., & Wolshon, B. (2013). Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C: Emerging Technologies*, *27*, 25–45. https://doi.org/10.1016/j.trc.2012.11.005

Na, H. S., & Grace, R. (2022). Influence of social networks and opportunities for social support on evacuation destination decision-making. *Safety Science*, *147*, 105564. https://doi.org/10.1016/j.ssci.2021.105564

Quarantelli, E. L. (1985). Social support systems: Some behavioral patterns in the context of mass evacuation activities. In B. J. Sowder (Ed.), *Disasters and mental health: Selected contemporary perspectives* (pp. 122–136). National Institute of Mental Health.

Siebeneck, L. K., & Cova, T. J. (2012). Spatial and Temporal Variation in Evacuee Risk Perception Throughout the Evacuation and Return-Entry Process. *Risk Analysis*, *32*(9), 1468–1480. https://doi.org/10.1111/j.1539-6924.2011.01781.x

Smith, S. K., & McCarty, C. (2009). Fleeing the storm(s): an examination of evacuation behavior during florida's 2004 hurricane season. *Demography*, *46*(1), 127–145. https://doi.org/10.1353/dem.0.0048

Whitehead, J. C., Edwards, B., Van Willigen, M., Maiolo, J. R., Wilson, K., & Smith, K. T. (2000). Heading for higher ground: factors affecting real and hypothetical hurricane evacuation behavior. *Environmental Hazards*, *2*(4), 133–142. https://doi.org/10.3763/ehaz.2000.0219

Wong, S., Shaheen, S., & Walker, J. (2018). *Understanding evacuee behavior: A case study of Hurricane Irma*. https://escholarship.org/uc/item/9370z127

Wu, H. C., Lindell, M. K., & Prater, C. S. (2012). Logistics of hurricane evacuation in Hurricanes Katrina and Rita. *Transportation Research Part F: Traffic Psychology and Behaviour*, *15*(4), 445–461. https://doi.org/10.1016/j.trf.2012.03.005

Yabe, T., & Ukkusuri, S. V. (2020). Effects of income inequality on evacuation, reentry and segregation after disasters. *Transportation Research Part D: Transport and Environment*, *82*, 102260. https://doi.org/10.1016/j.trd.2020.102260

Yang, K., Davidson, R. A., Blanton, B., Colle, B., Dresback, K., Kolar, R., Nozick, L. K., Trivedi, J., & Wachtendorf, T. (2019). Hurricane evacuations in the face of uncertainty: Use of integrated models to support robust, adaptive, and repeated decision-making. *International Journal of Disaster Risk Reduction*, *36*, 101093. https://doi.org/10.1016/j.ijdrr.2019.101093

Zeigler, D. J., Brunn, S. D., & Johnson Jr., J. H. (1981). Evacuation from a Nuclear Technological Disaster. *Geographical Review*, *71*(1), 1–16.

WUI-PEM: Wildfire Phased Zone Evacuation Methodology

Hui Xie
Movement Strategies (a GHD company), UK
Hui.Xie@ghd.com

Michael Kinsey
Movement Strategies (a GHD company), UK
Michael.Kinsey@ghd.com

Steve Gwynne
Movement Strategies (a GHD company), UK
Steve.Gwynne@ghd.com

Max Kinateder
National Research Council Canada, Canada
Max.Kinateder@nrc-cnrc.gc.ca

Paul Geoerg
Vereinigung zur Förderung des Deutschen Brandschutzes e.V., Germany
geoerg@vfdb.de

Roy Thun
GHD, USA
Roy.Thun@ghd.com

ABSTRACT

Wildfires are a growing threat to wildland-urban interface (WUI) communities. Globally over the last decade hundreds of wildfire related deaths have occurred along with hundreds of billions in economic losses. If the encroachment of fire into a community cannot be stopped, evacuation becomes necessary to save lives. In response to the increasing threat, WUI communities must properly prepare and implement an evacuation plan, as orchestrating an evacuation within a limited time frame is a complex challenge. While several fire and evacuation simulation tools have been developed to aid planning, a comprehensive methodology for constructing and assessing evacuation plans is still evolving. This article fills some of the current gaps with a simple six-step process to evaluate wildfire risk, individual evacuation zone traffic demand, and formulate a staged zone-based full evacuation plan. This methodology can be used for wildfire evacuation planning, evacuation plan management, and training.

Keywords:

Wildland urban interface, wildfire, evacuation planning, emergency management, simulation.

INTRODUCTION

A 'wildfire' can be defined as an unplanned vegetation fire produced by natural or human causes. Wildland urban interface (WUI) fires involve fire encroaching into populated communities causing loss of life and economic losses, including damage to property, infrastructure, and natural resources, as well ongoing concerns such as businesses and industries (Bento-Gonçalves et al., 2020; Schug et al., 2023). The development of the fire itself will be influenced by geography, weather patterns, and vegetation; the impact on affected buildings by construction materials/techniques and land use planning; while the capacity for communities to cope with an incident will be influenced by the resources available, the information available, and their ability to evacuate. These elements combine to produce the outcome in the context of planning effective mitigation and intervention efforts (i.e. emergency management and response).

Climate change has contributed to the increase of high-profile wildfires globally in recent years. For instance, extreme weather conditions contributed to the 2018 California wildfires, resulting in 100 deaths and an estimated $148.5 billion US dollars in economic losses, including damage to homes, businesses, and infrastructure (CAL Fire; Wang et al., 2021), while extreme heat promoted the 2019-2020 Australian bushfires resulting in an estimated $103 billion Australian dollars in losses and 33 deaths (Ogie et al., 2022). Over the decades, there is a tendency that more frequent and severe wildfires occur at many places around the globe, as illustrated in Figure 1. Figure 2 shows the tree loss by country using 2010 as a reference point. It is apparent that some countries have shown a dramatic loss, e.g. Portugal has a 39% loss, Sierra Leone a 33% loss, and Australia a 17% loss (Global

Forest Change). In response to the threat caused by the wildfire, it becomes a crucial requirement for all communities located at wildland urban interface to prepare for such a disastrous event. This, however, is challenging (Murray et al., 2023). As such events begin to encroach on occupied regions (only made more likely by the migration of people to the WUI), so the risk of exposure and the need for emergency response increases.

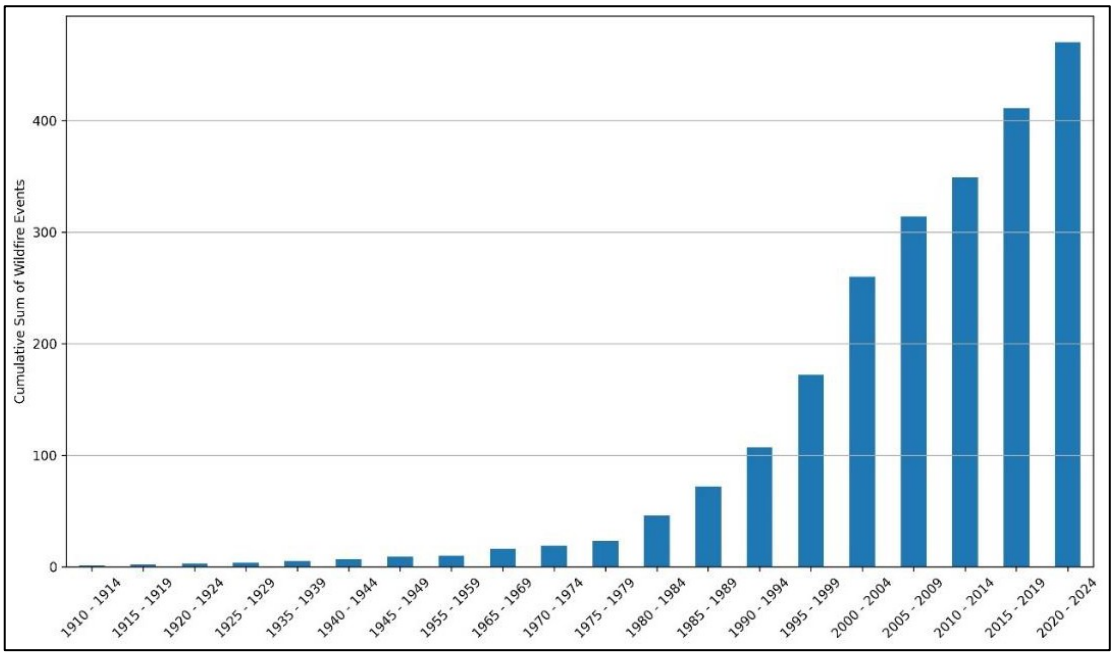


Figure 1: Occurrence of wildfire in five-year intervals. Data: (Global Forest Watch).

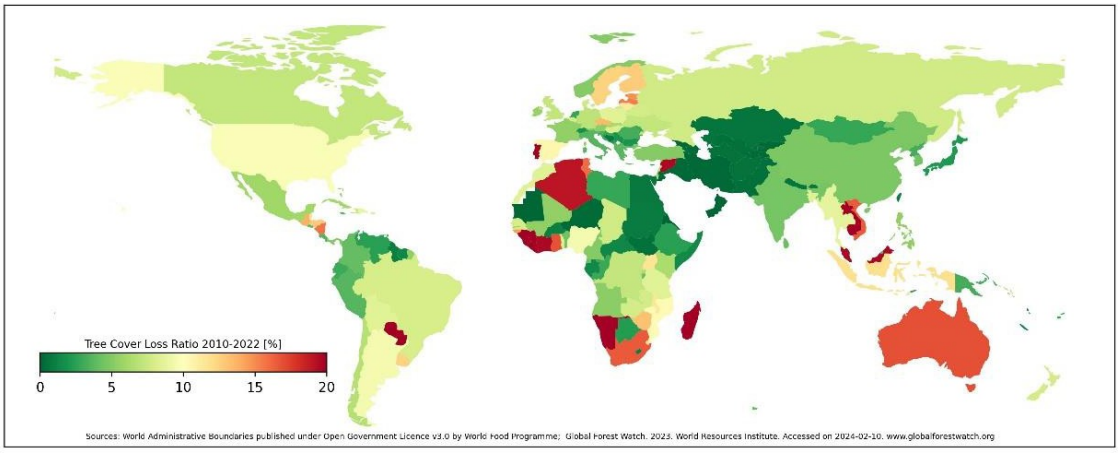


Figure 2. Tree loss due to wildfire between 2010-2022 using 2010 as reference point (Global Forest Watch).

The ability to carry out a large-scale evacuation is a key measure to safeguard communities when a fire threatens inhabited areas. It may be particularly important where mitigation measures have been ineffective and where real-time intervention cannot halt the progress of an oncoming fire front. Conducting such a large-scale evacuation requires a coordinated effort from multiple agencies and stakeholders, including emergency responders, local government officials, and community leaders to work together, taking into account all the relevant factors in evacuation planning and management. These factors include, but are not limited to, the nature of the fire threat (affected by the landscape, the vegetation, and the wind conditions affecting the movement of the fire front and therefore the time available for evacuation); the number, mobility and distribution of people to be evacuated; the availability of evacuation routes and target destinations; and the resources required for the evacuation (e.g., intervention, communication, vehicles and shelter etc.). Given the complex interaction between these factors and the importance of the evacuation performance (for both planning and emergency management), it is necessary to analyse the relationship between these factors.

- **The threat**: When a fire is detected, the direction and speed of fire spread, and its distance to the community combine to determine the fire risk, i.e., the level of threat posed to the community. The threat

also affects the range of the area to be evacuated and the available time for evacuation.

- **The evacuation demand**: The demographics, size, and distribution of the population within the community form the demand for evacuation. This demand is also affected by the layout and distribution of the community properties, including the residential streets within the community as well as the demographic and socioeconomic makeup of a community.
- **The routes for evacuation**: In most WUI communities, evacuation will need to be facilitated by the road network. That is, the evacuating population will need to leave the threated area via either familiar routes or designated evacuation routes through the road network to a place of safety. Thus, the layout and capacity of the road network, given their availability and use, will influence capacity and therefore the required evacuation time assuming demand is at a manageable level.
- **The resources required for the evacuation**: The resources required include fire intervention and traffic control, fire prevention and traffic control, the availability of vehicles for evacuation, and the notification system in place to convey the instructions to the residents (assuming that this affects the time to initiate evacuation along with preparatory activities).
- **The evacuation goal**: The availability of safe locations to accommodate the evacuees and their tools for transportation.

An effective evacuation plan for a community should be evidence-based, i.e. involve the assessment of threat level to different regions of the community (over time) and the effectiveness of different evacuation responses (e.g. including an attempt to balance the evacuation demand with the available routes and resources for evacuation), to increase the probability of a safe community evacuation.

The goal of developing this evacuation planning methodology is to take the above factors into account in planning or managing an evacuation, so that it can both minimise the risk of exposure to the fire threat, and stage the evacuation based on a zone-based division of the community, so that the resources are more effectively utilised and any plan put in place is more robustly tested. In other words, the evacuation plan should be designed to ensure that the risk of exposure to the fire threat is minimised, while also ensuring that the evacuation is carried out in a way that makes the best use of available resources. The plan is based on a zone division of the community given its layout, so that the evacuation can be designed, managed, and tested at the zonal level. It is recognized both that this methodology is draft in nature and broadly drawn given that the intention is that it might be generally applied. However, such methods are both uncommon and necessary both to encourage the appearance of evacuation models and structure their use.

QUANTIFYING PERFORMANCE

Ideally evacuation performance would account for all aspects of emergency management, including prevention, mitigation, preparedness, response, and recovery. Given the challenges posed in the collection and analyses of traditional outcomes and the significant variability of fire scenarios (e.g. making reliance on analytical derivations and empirical correlations much more challenging), wildfire evacuation modelling might be employed to quantify the interaction between the various elements and their consequences. Any model applied will inevitably contain multiple approximations and assumptions; however, coupled with subject matter expertise such efforts can reasonably assess the potential impact of certain factors, their interaction, and the sensitivity of the outcome to this interaction. That is, they might provide the means for communities, first responders, and emergency managers to define and explore certain ‘what-if’ scenarios. The use of modelling tools that consider traffic evacuation dynamics along with wildfire spread and pedestrian response might inform decision-makers about potential outcomes, vulnerabilities in their communities, and the most appropriate strategy to adopt (Ronchi et al., 2019; Kim et al., 2024).

There are a number of evacuation-related simulation platforms available (Ronchi et al., 2019; Intini et al., 2019), for instance, fire models (e.g. Spark, FARSITE, Prometheus, Phoenix, CAWFE, etc.), pedestrian models (e.g. EvacMod, FDS-Evac, Pathfinder, PedGo, Pedestrian Dynamics, JPS, etc.) and traffic models (e.g. OREMS, EVAQ, DYNA-SMART, AIMSUN, SUMO, TRANSIMS, VISSIM, etc.) that can be used to estimate evacuation outcomes. Most models simulate only one of the necessary features (i.e. fire, traffic, or pedestrian). There are several evacuation simulation tools available that attempt to capture the complexity of wildfire evacuation events – by including several aspects of the wildfire events including urbanEXODUS, Simtable, FLEET and WUI-NITY (Grajdura et al., 2022; Veeraswamy et al., 2018; Wahlqvist et al., 2021; Wildland Fire – Simtable; Fast Local Emergency Evacuation Times (FLEET); Genasys EVAC). These tools, amongst others that are emerging, would allow the practitioner to assess the evacuation of a community under different evacuation strategies and route availability – enabling the effectiveness and robustness of such strategies to be assessed. This is a pre-requisite for the methodology described below.

METHODOLOGY

The objective of the proposed evacuation planning methodology is to develop an effective evacuation plan for a threatened community with the aim for those individuals who need to evacuate, to do so safely and effectively. The method is based on a zone-based division of the region. By assessing fire risk and evaluating individual zone and full evacuation performance, the method provides guidance on constructing staged evacuation plans and applying evacuation simulations to select a plan based on performance metrics. It is important to note that this methodology is not tied to any specific modelling tool. However, it is essential to recognize that the development of this methodology imposes certain requirements on evacuation models. Specifically, these models must exhibit the flexibility necessary for effective evacuation planning and the capability to compare the performance of different evacuation plans using the proposed measures. Finally, this methodology can serve as training guidance in preparation for potential wildfire emergencies.

The methodology comprises six essential steps for analysing evacuation procedures and the metrics for selecting an effective plan. The first two are preparatory steps, which identify the geographical area for evacuation planning analysis and create a zone division of the area – the fundamental information for zone-based evacuation planning. Step 3 creates a simple estimate of exposure of different part of the area to the fire threat, based on the zone division. This step generates a rank of the zones with associated weight values, which will be used in the subsequent steps. This rank will be referenced for defining evacuation sequence of the full evacuation plans. And the weight values will be used in performance metric calculations for these plans, so that the fire risk is considered in identifying an effective plan. Step 4 defines three criteria for assessing the required safe evacuation times for the zones and Step 5 assesses the evacuation performance of individual zones through simulations according to these metrics. The outputs of these two steps are essential for determining the timing of the evacuation sequence in the last step. Step 6 (the last step) defines the components of staged full evacuation plans and suggests how to construct the plans using the outputs from previous steps. Step 6 utilises the results obtained from applying simulations to the plans and the proposed performance metrics to identify the final plan. These steps are outlined in detail below.

Step 1: Identify the location and range of the region to be considered for evacuation planning analysis.

The first step is to examine the target region that is susceptible to fire threats and identify the range of the area for evacuation planning or management analysis. Mapping platforms such as OpenStreetMap, Google Maps or Bing Maps can be used to examine the area, with consideration of the following criteria:

- The region should encompass the entire residential area and traffic network, inclusive of unimproved roads, relevant for evacuation analysis. Depending on the scale of evacuation planning, this could be a local community, a town, or even a city.
- The region should include evacuation objectives away from the threatened area. If the evacuation goal is not clearly defined, the region should cover a specific length (e.g., X kilometres) of the roads leading out of the populated area. Such a road segment defines a city exit within the region.

Step 2: Define evacuation group according to emergency management zone (EMZ).

For a large, populated community (e.g., a city), it might not be a credible strategy to order an evacuation of the whole population at the same time – especially if routes are limited and the community is densely populated. This is because the road network is seldom designed and developed with the aim to accommodate the surge of traffic demand from all residents. In practice, officials often divide such a large community into multiple emergency management zones (EMZs) so that the evacuation can be staged / managed at this level of resolution. While there are several factors (e.g., the population distribution, the layout of local road network) that can be considered in this process, the division of the area and the identification of the zone boundaries often follow political or geographical boundaries. This is primarily to facilitate the conveying of the message to people and their interpretation of the emergency management instructions. The development of this evacuation planning methodology is based on the zone division for the planning of emergency response – assuming the zone division is completed. The successive steps will assess the risk of exposure to the fire threat at different zones[1], and stage the evacuation of the zones to make an efficient use of the available road network capacity – assuming necessary resources such as notification and communication systems, and emergency management personals are in place to carry out the staged full evacuation plan.

[1] Zone and EMZ are interchangeable terms used throughout this paper.

Step 3: Assess fire risk to EMZs: Pre-Modelling Ranking.

To plan a staged evacuation for the EMZs, their exposure to the potential fire risk needs to be assessed first to prioritise their evacuation order. While there are many factors influencing the speed of the fire spread, range, intensity, and its subsequent impact on the community, here we propose a simple method to give a basic estimate of risk that can help structure the subsequence modelling. This simplified approach can be used when there is a lack of necessary information to perform a more detailed assessment of the fire risk (especially the spread of the fire itself). The approach should be enriched should both the information for fire modelling and fire simulation capability are available for fire risk assessment (Kalogeropoulos et al., 2023). This is beyond the scope of the description of this planning method.

When a fire is detected, or where there is a high risk of ignition at a nearby region with vegetation as fuel, the distance from the fire front to the EMZs can be used to assess the level of risk to which they are exposed. In general[2], the EMZs that are close to the fire, or the source of a potential fire should be ranked at the top in the evaluation order if the evacuation must be staged. To construct a fire risk factor for an EMZ that allows the comparison of the level of threat with the others, a normalised fire risk (FR_i) for the i^{th} EMZ based on the distance to a fire threat is defined as

$$FR_i(D_i) = \frac{\sum D_i}{D_i} \Big/ \sum \frac{\sum D_i}{D_i},$$

where D_i is the shortest distance from the fire front to the perimeter of each EMZ.

The method of assessing the fire risk to EMZs is illustrated using a hypothetical example shown in Figure 3. The community consists of four EMZs labelled from A to D. It is assumed that a fire starts at the top left corner of the region with a red arrow showing the possible propagation direction towards the community. Subject to the distance to the location of fire threat and the direction and speed of potential fire spread, the EMZs may be exposed to different levels of risk. In general, the closer to the fire front a community is when the fire propagation is detected, the higher the fire risk of the place will be. Thus, people living in this region should be evacuated with priority, and the distance can be used to construct a risk indicator (although other factors, such as weather and wind conditions or the geometry of the road network certainly also affect the potential risk at a given location).

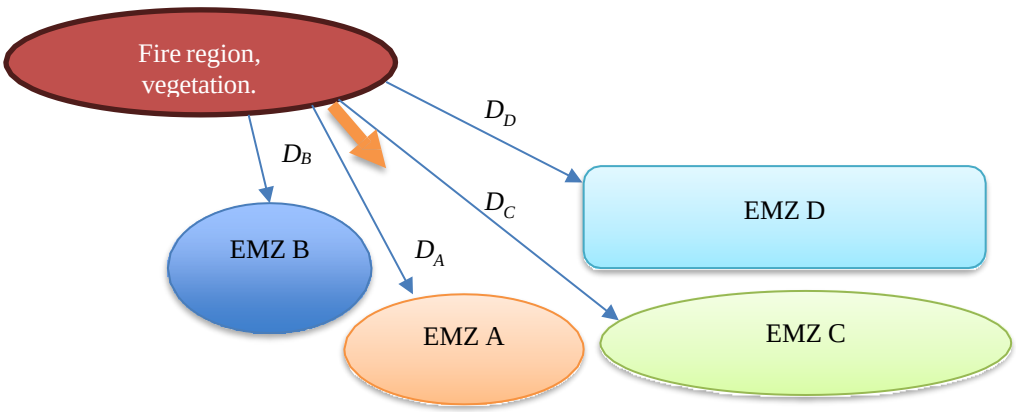


Figure 3. The relative location of EMZs and the fire.

Table 1 lists the values of D_i and the calculated value of FR_i for the four EMZs in the example. The closer an EMZ is to the fire, the higher the fire risk indicator is. For instance, EMZ C is four times further away to the fire than EMZ B, the risk value of EMZ B is about four times of that of EMZ C. Note that FR_i is a signal-factor risk indicator; it does not include the other potential influencing factors in the assessment of the fire risk to the EMZs. For example, equivalent vegetation type and distribution between the fire front and each region is assumed. Additionally, to be fair with the comparison of the fire risk using FR_i, it is also assumed that the EMZs are comparable in size, i.e., none of the EMZs is significantly larger or smaller than the others. If there is any EMZ that is significantly smaller than the others, it may be merged with a nearby EMZ which has similar D_i for planning analysis.

[2] For example, this might not be the case should there be a change of wind direction, or where there are measures to separate or protect the zones, so that some zones may not catch fire. Again, we are not trying to cover all the possibilities but the general concept.

Table 1. The calculation of fire risk based on distance to the source of fire risk.

	EMZ A	EMZ B	EMZ C	EMZ D
D_i (km)	4	1.5	6	4
FR_i	0.19	0.50	0.13	0.19

Step 4: Assess the evacuation performance of individual EMZs: Evacuation Assuming Ideal Management Conditions.

To create an effective staged and zone-based evacuation plan, the traffic demand from each zone and the expected duration of use are two key pieces of information for planning the evacuation and optimising route usage. To be able to estimate the information, this step defines three exit stages from their start location (households) to destination (final place of safety), which set the criteria for estimating the times for people to escape from the danger. Based on this definition, Step 5 introduces the method to estimate the information through simulations.

Unlike the traditional building evacuation planning, which normally includes a signal stage of evacuation to a place of safety, the evacuation route from residential areas to a final place of safety can be broadly divided into three sections, which correspond to three exit stages of the evacuation (see Figure 4 and an example illustrated in Figure 5). The first section includes roadways (primarily for local traffic) within EMZs connecting households and the nearest arteries in the city, which is normally located at the boundary of the corresponding EMZ. The location where the perimeter of the EMZs intersects with the main roads can be defined as **zone exit**. During this stage of evacuation, people who have decided to respond emerge from their homes, join the evacuation traffic follow on local roadways towards the main roads (i.e., zone exit). Progress will have been made once they have arrived at these main roads; however, they still need to continue driving along the main roads towards the boundary of the city and then the place of safety. They are still vulnerable, but less so than when they had first initiated movement.

The second section of the evacuation route is from the main roads within the city towards the city boundary. During this stage of evacuation, there will be merging flows from the same or other EMZs on the main roads. Therefore, people might not be able to drive at their desired speeds and may also experience congestion – depending on the population size/density, demand levels, route capacity and the conditions faced. The location where the perimeter of the city intersects with the main roads out of the city can be defined as the **city exit** (see Figure 5). When people pass the city exits, it usually means they have left the populated area. Therefore, less merging traffic from side roads will be expected, and people should therefore be able to drive at relatively higher speeds as they move further away from the city.

The last section of the evacuation route is normally from major arteries or freeway leaving the city towards the remote **regional exit** (see Figure 5) representing a place of safety, such as planned sheltering places or a remote location far away from the wildfire threat. If there are no planned sheltering places and people are heading towards their own elected destinations once they leave the city, then a location on the freeway that is relatively far away from the area being threatened can be defined as a regional exit for the modelling purpose.

In this step, it is required to perform a static analysis of the zone map including the traffic network. The goal is to identify the three types of exits and determine the evacuation route options to leave through the zone and city exits towards the final reginal exits. When planning the evacuation route for each zone, the consideration for route choice includes nearest, fastest (and route familiarity if the information is available) route towards assigned sheltering places. The planning should also consider the likely use of the roads by people in other zones. The insights gained from Step 4 serve as input for the subsequent evacuation performance analysis in the next two steps.

Step 5: Stress-Testing: individual zone evacuation analysis.

Before complex situations during a full evacuation (for instance, as a result of management or intervention) are assessed, it is essential to develop an understanding of the evacuation performance of individual zones. In essence, how each zone might perform should other zones be successfully delayed in their response. This step conducts a stress test of the response from each zone, given the routes they use, to inform the development of city-wide evacuation plan and the test outlined in Step 6. In the stress test, each zone is evacuated alone with instant response (to encourage the development of peak levels of congestion) and the route choice obtained in Step 4 under two traffic conditions:

1. **Zero background traffic** – representing the assumption that people in the EMZ have priority in the order of evacuation and the attempt to manage the background traffic is successful.
2. **Certain amount of background traffic** - representing the assumption that either (a) the attempt to manage the background traffic is partially successful or (b) the potential for interacting with evacuees

from other zones given they ignored previous instructions (here represented implicitly by a reduction in available route capacity).

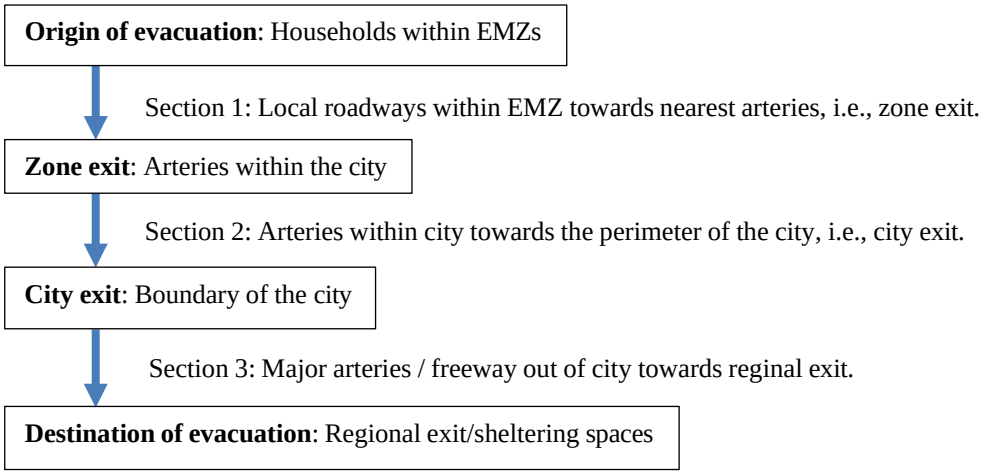


Figure 4. Three stages of the wildfire evacuation.

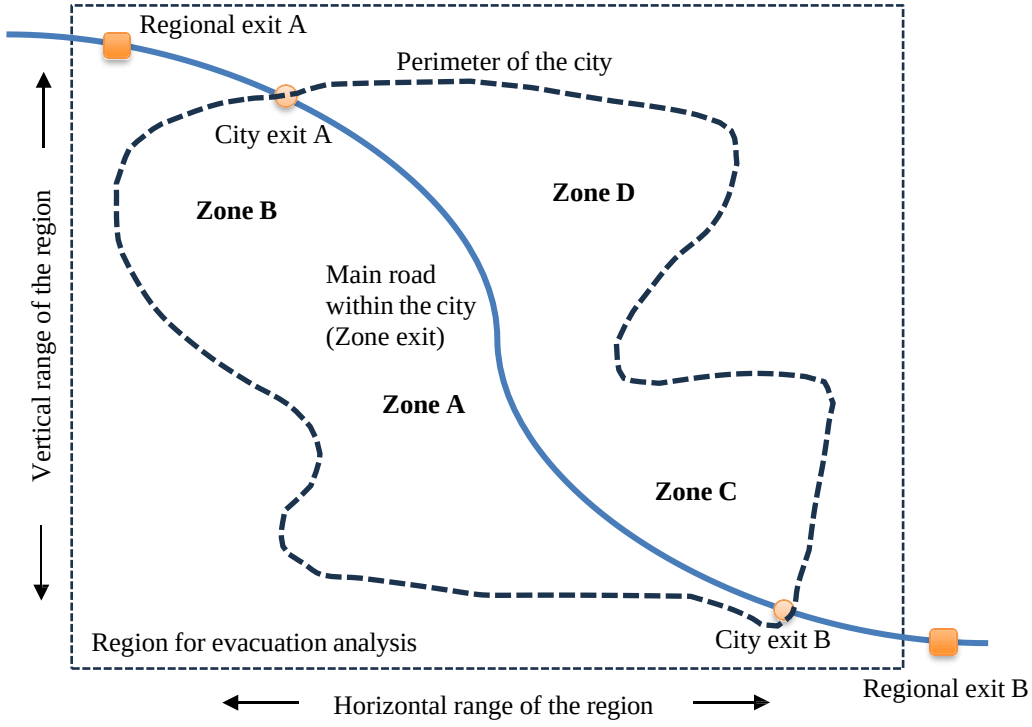


Figure 5. The range of the region for evacuation analysis and three types of exits.

Under the stress testing condition (i.e., instantaneous response / focus on response of individual zones), the simulation results produced from the first test traffic condition give an indication of the minimum time required to move the sub-population to safety, while the results produced from the second test traffic condition indicate a more representative estimate of the time required for the sub-population to evacuate the zone in general. The first time is useful for planning during an emergency in which a particular zone is facing an immediate threat especially in establishing locations of potential route congestion, while the second time is useful for making a city-wide evacuation plan reflecting more representative clearance times for each sub-population. Both might be used to prioritize zone evacuation and suggest design modifications to reduce sources of traffic congestion.

Corresponding to the three stages of evacuation outlined in Step 4, the evacuation performance of each zone has three indicators:

- Ez_{i0} - the time to clear the i^{th} zone alone.
- Ec_{i0} - the time for the subpopulation of the i^{th} zone alone to clear the city exits.

- Er_{i0} - the time for the subpopulation of the i^{th} zone alone to reach the regional exits.

Given these indicators are obtained without considering the interaction with the subpopulations in other zones, these form the baseline for comparison with the evacuation performance obtained from all EMZs evacuation analysis in next step. These indicators are also the bare minimum requirements for safe evacuation of the zones.

Step 6: All EMZs evacuation analysis in response to various potential fire threats.

The final step of developing a city-wide evacuation plan involves determining the order and timing of evacuations for the EMZs. This process considers the factors examined in all previous steps, including the zone divisions (Step 2) within the examined area (Step 1), fire risk assessment for each zone (Step 3), the three evacuation stages (Step 4), and the minimum requirement derived from stress-testing of individual zone (Step 5). A general approach for constructing effective evacuation plans involves:

- **EMZ evacuation sequence**: Define the EMZs evacuation sequence, with reference to the order of the fire risk from high to low, to stage the evacuation that allows those who are closer to the fire risk to leave with priority.
- **Timing Considerations:** Define the timing of the evacuation sequence, with reference to the times to clear the actively evacuating zone(s), and their times to clear city exits. The goal is to give the evacuating zone(s) sufficient time to clear while avoiding overloading the road network with traffic from multiple zones simultaneously or within a short period of time.
- **Route Planning**: Plan the use of route for the EMZs with reference to the availability of the regional exits (i.e., the location and capacity of the evacuation target), the road network connectivity and capacity, and the resources that can be deployed to manage the traffic. The goal is to efficiently utilise the capacity of the road network, while minimising the interaction between evacuation traffic from different zones to prevent bottlenecks and delays. The planning conducted in this step considers the collective traffic demand from multiple zones that may need evacuation within a limited time window. This differs from the route planning in Step 4, which focuses on individual zone traffic demand.

When considering various combinations of the first two components in the planning process (assuming an optimised route planning or a consistent basis), there are two extreme base plans that impact road capacity differently. The first one is to evacuate all EMZs simultaneously (i.e., without any delay gaps). The second one is to evacuate all EMZs strictly in sequence, with gaps set to each zone's clearance time down the line. The first one, which is equivalent to a stress test, exerts maximum pressure on the traffic network and may lead to significant congestion and delays. The second one, which minimises pressure on the traffic network, underutilises the road network's capacity, resulting in unnecessarily extended total evacuation time – in an attempt to manage road conditions. Therefore, these two base plans serve as boundaries for constructing other feasible evacuation strategies. The goal is then to strike an ideal balance between the planning components, while ensuring the safe and efficient evacuation of all residents.

Those feasible plans that strike a balance between evacuation efficiency and pressure on the traffic network involve intricate interaction among evacuating people from different zones. These evacuation plans can be assessed through computer models that represent corresponding evacuation scenarios. However, evaluating these plans based solely on total evacuation time does not fully address safety concerns. For instance, two plans - one accounting for fire risk to decide evacuation sequence and another ignoring it, might yield similar evacuation times. To address this, we propose a composite evacuation time, defined as the sum of the three zone-specific indicators corresponding to the three stages of evacuation, for each plan. This metric, coupled by the fire risk factors for the zones obtained in Step 3, is then used in the assessment of the simulation results to evaluate the effectiveness of evacuation plans.

Assuming a set of evacuation plans have been defined as P_j with the combination of evacuation sequence and timing of evacuations for the EMZs (and a consistent route planning). The performance metrics defined as composite evacuation times weighted by the fire risk factor for these plans are as follows:

- $\sum_i Ez_i(P_j) * FR_i$ - the composite evacuation time to clear the zones,
- $\sum_i Ec_i(P_j) * FR_i$ - the composite evacuation time to clear the city exits,
- $\sum_i Er_i(P_j) * FR_i$ - the composite evacuation time to reach the regional exits,

where $Ez_i(P_j)$, $Ec_i(P_j)$ and $Er_i(P_j)$ are the zone exit times, city exit times and reginal exit times of the corresponding zones within each of the plans. These exit times can be obtained through computer simulations that simulate the evacuation plans for the region.

Finally, the metrics are compared against the baseline defined as the composite evacuation time obtained through

individual zone evacuation analysis performed in Step 5 (as shown in Figure 6). The deviations (dz_j, dc_j and dr_j) from the baseline serve as reference for assessing plan effectiveness. The ultimate objective is to identify a plan that minimises deviations while ensuring safe evacuation. For instance, if the priority is to evacuate people from their zones to the main road first, then focus on minimising dz_j when choosing a plan. If the safe evacuation goal also involves clearing city exits and reaching reginal exits, consider minimising dc_j and dr_j as well in the decision-making process.

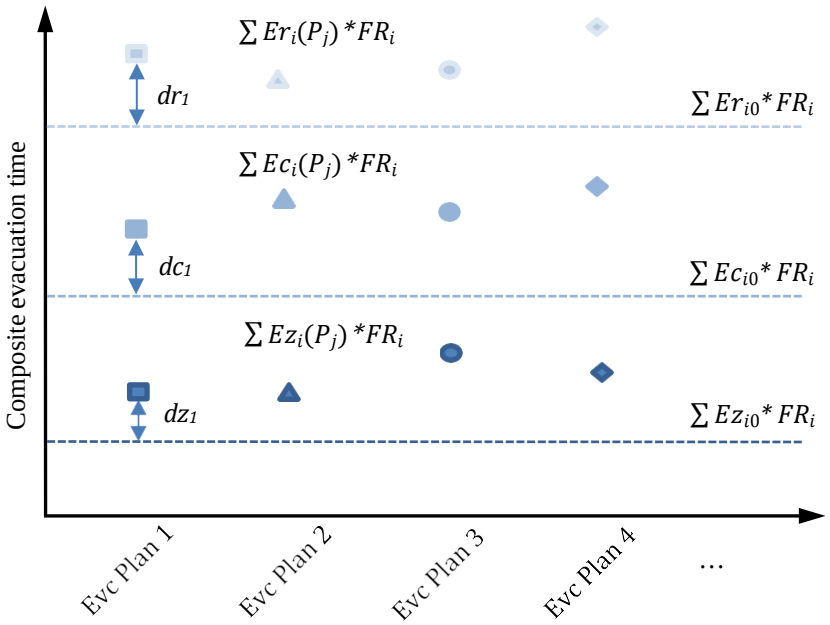


Figure 6. The performance metrics for evaluating the evacuation plans.

In Step 3 of the analysis, one fire condition is assumed. If there are multiple potential fire threats to be considered for the region, the above six-step evacuation planning analysis should be conducted for each of them. The analysis of the fire threat should also include the assessment of its impact on the availability of city exits, and regional escape routes. This information feed directly into route planning in the Step 5 and Step 6 analyses.

The above six-step evacuation planning considers the static status of the fire threat, i.e., the fire risk assessment is only based on its distance to the EMZs in the region. When a final plan is identified from the planning analysis, it should be compared against the projected development of the fire threat to assess the emergency level to further adjust the plan if necessary. The emergency level (EL) can be assessed through

$$EL = \frac{T_t * (1 + S_m)}{T_f}$$

where T_t is the total evacuation time from simulating the identified evacuation plan, S_m is a safety margin and T_f is the projected time for the fire to reach the perimeter of the community. T_f can be obtained through fire simulation, considering vegetation location, weather conditions and wind direction (i.e. direction of approach) etc.

If $EL < 1$, the emergency level is manageable since either the fire front is still far away, or it takes longer time than the required safe evacuation time of the community. The situation can be monitored to decide whether and when to issue an evacuation order.

If $EL \geq 1$, the emergency level is high and the fire can potentially be harmful to the evacuating people, especially to those EMZs ranked high in the fire risk assessment. An immediate response including evacuation should be considered.

LIMITATIONS AND DISCUSSIONS

This paper presents a planning methodology to develop evidence based and quantifiable evacuation strategies for WUI communities threatened by wildfires. This approach is based on the idea that computer models and calculations can be used to describe evacuation performance. As any model, the work presented here, is a

simplified representation of the complexities that can be expected in a 'real' incident and it should be noted that the approach selected has certain limitations; for example, it is proposed that Fire Risk can be represented in terms of Euclidian distance from an encroaching fire; however, other metrics (e.g., estimated travel times) might be equally or more representative. Further, how quickly a community needs to issue an evacuation order is determined by multiple factors (e.g., geography, vegetation, weather, wind, community layout, means of transportation, etc.) (Kalogeropoulos et al., 2023), the evacuation process involves heterogenous responses to evacuation alerts, and associated behavioural itineraries (Vaiciulyte et al., 2021). The inclusion of these two and other potential factors is a relatively minor task to complete and is left for future method development. The method shown is therefore presented to show a general application of evacuation model output and the types of insights that might be provided.

CONCLUSION

The evacuation of areas affected by wildfires is likely to become more commonplace and involved an increasing number of communities. Wildfire evacuation pose numerous challenges due to multiple factors and inherent uncertainties involved. Given this, the emergency plans might become more complex and sensitive to increasing numbers of variables. Coupled with the challenges with relying on historical data, the capacity to quantify evacuation performance will be increasingly important – to enhance design and improve between alternative community designs and emergency plans. Despite the existence of various evacuation models tailored for different scenarios, there remains a gap in applying these tools to plan effective response measures. To address this issue, this article has outlined a simple means by which such a quantitative assessment might be structured – a methodology that can be employed to further develop evacuation models and establish common performance metrics. As a work in progress, this method is still in development. We anticipate both further enhancement of the method with practical details and application scenarios that will demonstrate its applicability with the proposed measures.

REFERENCES

Bento-Gonçalves, & A., Vieira, A. (2020). Wildfires in the wildland-urban interface: Key concepts and evaluation methodologies, *Science of The Total Environment*, vol. 707, 135592, ISSN 0048-9697. https://doi.org/10.1016/j.scitotenv.2019.135592

CAL Fire, *California 2018 Incident Archive*, California Department of Forestry and Fire Protection. Retrieved February 14, 2024, from https://www.fire.ca.gov/incidents/2018

Fast Local Emergency Evacuation Times (FLEET), VMASC, Old Dominion University, Retrieved February 14, 2024, from https://fleet.vmasc.odu.edu/

Genasys EVAC | Genasys Protect platform (Formerly Zonehaven). Retrieved February 14, 2024, from https://genasys.com/genasys-protect-evac/

Global Forest Change, Global Land Analysis & Discovery, Retrieved February 10, 2024, from https://glad.earthengine.app/view/global-forest-change

Global Forest Watch. World Resources Institute. Retrieved February 10, 2024, from https://www.globalforestwatch.org

Grajdura, S., Borjigin, S., & Niemeier D. (2022). Fast-moving dire wildfire evacuation simulation, *Transportation Research Part D: Transport and Environment*, Volume 104, 103190, ISSN 1361-9209, https://doi.org/10.1016/j.trd.2022.103190

Intini, P., Ronchi, E., Gwynne S. M. V., & Pel A. (2019). Traffic Modeling for Wildland–Urban Interface Fire Evacuation, *Journal of Transportation Engineering, Part A: Systems*, Volume 145, Issue 3, https://doi.org/10.1061/JTEPBS.0000221

Kalogeropoulos, N., Mitchell, H., Ronchi, E., Gwynne, S. M. V., & Rein, G. (2023), Design of stochastic trigger boundaries for rural communities evacuating from a wildfire, *Fire Safety Journal*, vol. 140, 103854, https://doi.org/10.1016/j.firesaf.2023.103854

Kim, A. M., Beverly, J. L., & Zahid A. A. (2024). Directional analysis of community wildfire evacuation capabilities, *Safety Science*, vol. 171, 106378, ISSN 0925-7535, https://doi.org/10.1016/j.ssci.2023.106378

Murray, A. T., Baik, J., Figueroa, V. E., & Rini, D. et al. (2023). Developing effective wildfire risk mitigation plans for the wildland urban interface, *International Journal of Applied Earth Observation and Geoinformation*, vol. 124, 103531, ISSN 1569-8432, https://doi.org/10.1016/j.jag.2023.103531

Ogie, R., Moore, A., & Wickramasuriya, R. et al. (2022). Twitter data from the 2019–20 Australian bushfires reveals participatory and temporal variations in social media use for disaster recovery. *Scientific Reports*, volume 12, Article number: 16914. https://doi.org/10.1038/s41598-022-21265-6

Ronchi, E., Gwynne, S. M. V., Rein, G., Intini, P., & Wadhwani, R. (2019). An open multi-physics framework for modelling wildland-urban interface fire evacuations, *Safety Science*, vol. 118, 868–880, ISSN 0925-7535, https://doi.org/10.1016/j.ssci.2019.06.009

Schug, F., Bar-Massada, A., Carlson, & A.R. et al (2023), The global wildland–urban interface, *Nature*, vol. 621, 94-99. https://doi.org/10.1038/s41586-023-06320-0

Vaiciulyte, S., Hulse, L. M., Veeraswamy, A., & Galea, E. R. (2021). Cross-cultural comparison of behavioural itinerary actions and times in wildfire evacuations. *Safety Science*, vol. 135, 105122. https://doi.org/10.1016/j.ssci.2020.105122

Veeraswamy, A., Galea, E. R., Filippidis, L., & Lawrence, P. J. et al. (2018). The simulation of urban-scale evacuation scenarios with application to the Swinley forest fire, *Safety Science*, vol. 102, 178-193, ISSN 0925-7535, https://doi.org/10.1016/j.ssci.2017.07.015

Wahlqvist, J., Ronchi, E., Gwynne, S. M. V., Kinateder, M., & Rein, G. et al. (2021). The simulation of wildland-urban interface fire evacuation: The WUI-NITY platform, *Safety Science*, Volume 136, 105145, ISSN 0925-7535, https://doi.org/10.1016/j.ssci.2020.105145

Wang, D., Guan, D., & Zhu, S. et al. (2021). Economic footprint of California wildfires in 2018. *Nature Sustainability*, *4*, 252–260. https://doi.org/10.1038/s41893-020-00646-7

Wildland Fire – Simtable. Retrieved February 14, 2024, from https://www.simtable.com/emergency-management/fire/

It's a Machine Learning world: no future for Complex Event Processing?

Anne-Marie Barthe-Delanoë
IMT Mines Albi, University of Toulouse
anne-marie.barthe@mines-albi.fr

ABSTRACT

In 2010, Complex Event Processing was a groundbreaking approach for data processing for Early Warning Systems and Decision Support Systems in the crisis management domain. Although Machine Learning has eclipsed CEP in the ISCRAM community these last few years, these two approaches do not fulfil the same objectives. With this article, a quantitative analysis is made on papers related to these domains to understand the current trends and identify the next challenges the ISCRAM community should address. Combining CEP with ML can help to overcome the strong assumptions made towards the CEP rules this last decade, such as the availability of experts, structured data, etc. It is then recommended that the ISCRAM community explore the hybridisation of these approaches to improve the responsiveness in real time of EWS and DSS in crisis situations, emphasising the automatic, non-supervised generation of CEP rules.

Keywords

Complex Event Processing, Machine Learning, Early Warning, Decision Making, Trends.

INTRODUCTION

Early Warning Systems (EWS) (UNDRR, 2007) and Decision Support Systems (DSS) rely on data analysis and processing. In the context of EWS, it is crucial to identify, analyse and predict potential hazards and their outcomes. Similarly, in the context of DSS, valuable information must be extracted to support decision-making. Complex Event Processing (CEP) was a top trending event processing approach for data input for EWS and DSS under time-sensitive applications in the last decade, from financial fraud monitoring to crisis management. However, in crisis management, CEP now seems overshadowed, not to say replaced, by Machine Learning (ML) and all the artificial intelligence (AI) related disciplines despite its numerous strengths. For example, the last publications on CEP applications for EWS and DSS in the ISCRAM digital library were published a decade ago (Barthe-Delanoë et al., 2014; Kovordanyi et al., 2014; Pottebaum et al., 2011).

But is this a real trend or just a side effect of all the buzz about AI in the latest research projects and academic publications? Are these approaches exclusive or complementary for real-time data analytics for EWS and DSS in crisis management? Is there a future for CEP in this ML world? The research question is then: What are the current uses of CEP and ML to feed EWS and DSS, from a data analytics perspective?

To answer these questions, this position paper explores the relationship between CEP and ML approaches for data analytics in the context of information systems for crisis management. This paper will study the synergies, complementarities, and potential challenges of integrating ML and CEP for more effective EWS and DSS for crisis management. It will also focus on how the ISCRAM community deals with it now and which challenges they might address in the future.

This paper is structured as follows. First, a brief reminder of the characteristics of each approach is made. The second part will analyse the state-of-the-art in a significant academic publication database. The main takeaways of the analysis will then be addressed to the ISCRAM community in a third part to spark future discussions about efficiency in mining and analysing multiple flows of data in real-time to feed EWS and DSS, before concluding with insights for future research topics for the ISCRAM community.

WiP Paper – Open Track
Proceedings of the 21st ISCRAM Conference – Münster, Germany May 2024
Berthold Penkert, Bernd Hellingrath, Monika Rode, Adam Widera, Michael Middelhoff, Kees Boersma, Matthias Kalthöner, eds.

DATA PROCESSING AND ANALYTICS

DSS and EWS are based on information from the field (real-time data), and historical information (capitalised knowledge). The need to feed these systems with data is facing many constraints: real-time data processing, historical data processing, and management of numerous data sources, sometimes competing, sometimes uncertain, and sometimes temporarily unavailable. The data may also be available in the form of a continuous, ordered flow over time (data stream) or in the form of a cloud of data that is not necessarily ordered. It will also be necessary to aggregate low-level data (e.g. calculating an average temperature over the last 5 minutes of collected data) or to match patterns (e.g. A happening before B and after C within a 10-hour time window). This allows for extracting the necessary information at the right level (regarding the stakeholders' needs) to be actioned for risk assessment, monitoring, forecasting, or any other purpose related to early warning or decision support. Two approaches are mainly used to achieve these data processing and analytics steps in a time-sensitive context: Complex Event Processing (CEP) and Machine Learning (ML).

Complex Event Processing

Event processing is a method of monitoring and analysing (processing) information flows (data) about things that happen (events) to draw a conclusion from them. Complex Event Processing (CEP) is a set of concepts and techniques for processing events in real-time and extracting information from event streams as they arrive (Etzion & Niblett, 2010).

The objective of CEP is to identify significant events (such as opportunities or threats) in real-time situations and respond to them as quickly as possible. These significant events are complex because they cannot be detected or observed directly; they must be identified by deduction, analysis and correlation of observed elementary events (Luckham, 2011). For example, the combination of events embedding the temperature measure of sensors located in various areas of a building can lead to emitting an early fire alert event, indicating an incipient fire as the temperatures are slowly increasing with no abrupt temperature changes. The low latency of real-time data processing (under 10 µs in the state-of-the-art CEP engines) combined with the ability to face high volumes of data makes CEP a top technology for the Internet of Things community (Lima et al., 2022; Shah et al., 2022).

On the technical side (see Figure 1), the CEP system relies on manually pre-defined business rules (implemented as event patterns, event filters, event aggregators, causality rules, timing rules, etc.) continuously executed by the CEP engine over the incoming data streams of events that feed the event cloud. If a rule is verified, then the resulting complex event is generated. The event streams are collected through subscriptions to various heterogeneous event sources: sensors, log files, social networks, reports, etc. The generated complex events can also be collected and processed to deduce other complex events.

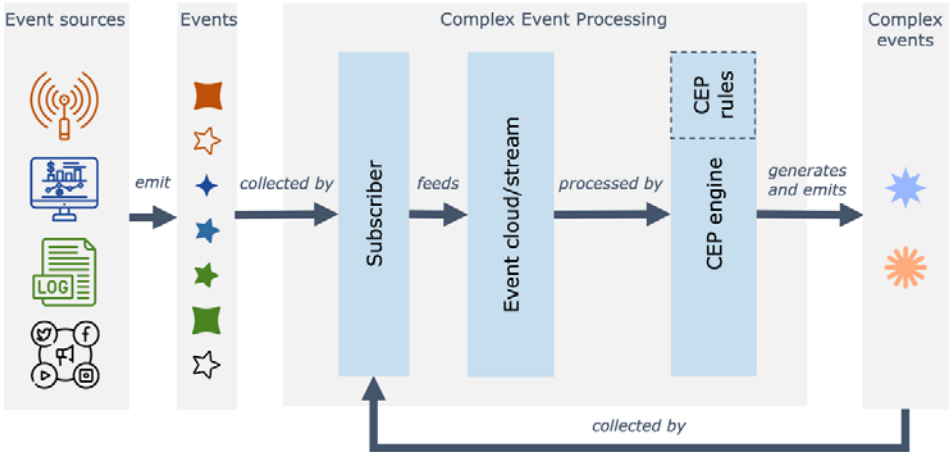


Figure 1. Principles of Complex Event Processing

It must be noted that the time dimension is natively part of the CEP system, allowing it to handle time frames (fixed or sliding) and the non-occurrence of an event. Moreover, CEP is not yet another expert system. CEP is not about decision-making as a standalone solution but about creating knowledge about a system in real-time (Fülöp et al., 2012). The generated complex events are consumed by EWS and DSS applications to feed their knowledge base, Common Operational Picture, and any component relying on situational awareness in a crisis management context, as shown by (Barthe-Delanoë et al., 2014; Kovordanyi et al., 2014; Pottebaum et al., 2011).

Machine Learning

Machine Learning (ML) was defined by (Samuel, 1959) and later by (Mitchell, 1997) as the field of study that enables computers to learn and improve from experience without being explicitly programmed. ML is a branch of Artificial Intelligence (AI) that aims to find the underlying relationships (i.e., models) between data. ML algorithms can analyse large volumes of data to identify patterns and trends that may raise potential opportunities and threats: detecting anomalies in a financial system, predicting failures in aircraft maintenance, identifying structural threats in a building, etc.

Technically, ML algorithms are trained, tested, and evaluated on historical data, split into a training set, a test set, and an evaluation set (see Figure 2). The outcome is a model called a predictor, which is then fed with new data and produces a prediction. The nature of the prediction depends on the type of the trained ML algorithm. It can be a discrete value, a series, a classification, a clustering, a probability, an association, etc.

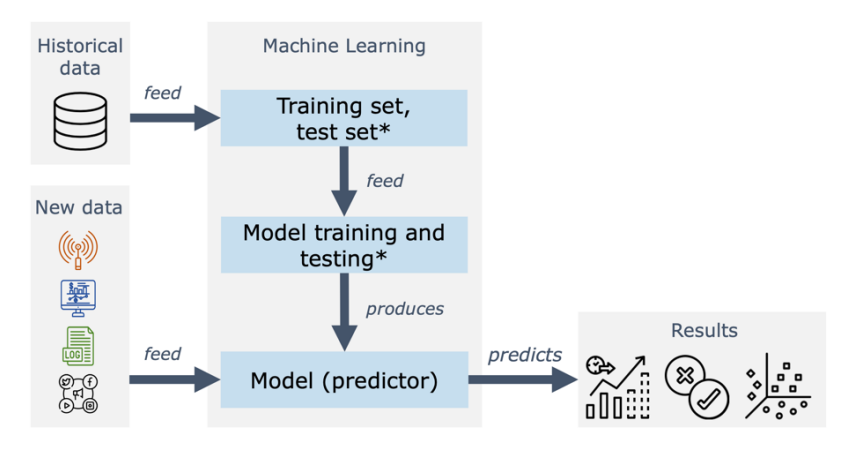


Figure 2. Principles of Machine Learning
(*evaluation step is often included with the test step)

ML can handle complex data better than classical data analytics approaches: datasets with many features and individuals; unstructured data (text, images, etc.); non-linear relationships among features. It can also make predictions, e.g. financial crisis prediction (Jaquart et al., 2021) and identification of disaster news using Natural Language Processing techniques (Chasseray et al., 2021; Domala et al., 2020).

Application to DSS and EWS

In EWS and DSS applications for crisis management, it is necessary to ensure the following requirements regarding the data analytics part: real-time detection of events and patterns of events, forecasting and prediction of the situation, low latency for timely decision-making, and scalability to handle large volumes of data.

Given the above characteristics of each of the CEP and ML approaches, a synthesis of their advantages and limitations is provided in Table 1.

Table 1. Advantages and limitations of CEP and ML for EWS and DSS (crisis management context)

Characteristics	CEP	ML
Real-time detection of events	Yes (real-time flow of data)	No
Real-time detection of patterns	Yes (use of predefined rules)	No
Forecasting and prediction	No	Yes (training on historical data)
Low latency	Yes (<10 µs)	Model dependent
Scalability	Yes (Big Data context)	Yes (Big Data context)

CEP is natively designed to detect events in real-time and complex patterns in huge data streams with low latency, making this approach suitable for time-sensitive applications. On the other hand, ML algorithms can uncover hidden patterns in historical data and make predictions about future events. They can also natively handle large volumes of data.

CEP, ML: WHERE ARE WE?

Overview of the Publication Trends

The first step of the inquiry about the position of CEP and ML in the literature is exploratory data analysis on the publications retrieved through six SCOPUS[1] queries, as shown in Table 2. These queries look for the presence of the CEP and ML approaches with a focus set on crisis management (queries *R1* and *R2*, respectively), on the DSS and EWS (queries *R3* and *R4*, respectively), and finally, on DSS and EWS for crisis management (queries *R5* and *R6*, respectively). The subject areas of all the queries were limited to the relevant ones, given the technical topic and the application domain (*dom*): computer science, engineering, mathematics, decision sciences, social sciences, business, management and accounting, energy, economics and finance.

Table 2. Set of SCOPUS queries

Req.	CEP	Req.	ML
R1	TITLE-ABS-KEY("event process*" AND ("cris* manag*" OR "emer* manag*"))	R2	TITLE-ABS-KEY ("machine learn*" AND ("cris* manag*" OR "emer* manag*"))
R3	TITLE-ABS-KEY("event process*" AND ("decision support*" OR "early warning*") AND system*)	R4	TITLE-ABS-KEY ("machine learn*" AND ("decision support*" OR "early warning*") AND system*)
R5	TITLE-ABS-KEY("event process*" AND ("decision support*" OR "early warning*") AND system* AND ("cris* manag*" OR "emer* manag*"))	R6	TITLE-ABS-KEY("machine learn*" AND ("decision support*" OR "early warning*") AND system* AND ("cris* manag*" OR "emer* manag*"))
dom	AND (LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "ENGI") OR LIMIT-TO (SUBJAREA , "MATH") OR LIMIT-TO (SUBJAREA , "DECI") OR LIMIT-TO (SUBJAREA , "SOCI") OR LIMIT-TO (SUBJAREA , "BUSI") OR LIMIT-TO (SUBJAREA , "ENER") OR LIMIT-TO (SUBJAREA , "ECON"))		

From the raw results of these six queries[2] (before the screening of each paper), it can be observed (Table 3) that the literature is more extensive regarding ML than CEP, which is explained by the variety of techniques and tools that can be found under the umbrella of ML compared to CEP.

Table 3. Raw results of the SCOPUS queries

Req.	CEP	Req.	ML
R1	31 documents	R2	329 documents
R3	114 documents	R4	6,984 documents
R5	10 documents	R6	24 documents

The gathered results are then plotted (Figure 3) to observe the overall distribution of publications by year for each 2-uple of queries: R1 vs R2, R3 vs R4, and R5 vs R6. It must be noted that since 2024 is at its beginning (to date of this review), the 2024 figures cannot be interpreted as for the previous years. Also, some publications from late 2023 are not yet referenced. A few duplicates were found between queries results (Table 4).

In Figure 3, a first observation is that the distribution of CEP-related publications peaks in the early 2010s, then decreases and reaches an asymptote (between 1 or 2 publications per year in the 2020s). On the other end, the ML-related publications tend to increase over time since mid-2010s with an exponential growth in the 2020s. It appears that the naïve impression of CEP being overshadowed by ML in the academic literature is correct from a quantitative point of view, especially in the case of ML/CEP for DSS and EWS applications, without connection to a specific business domain. However, if the focus is set on DSS and EWS restricted to crisis management applications, the number of publications drops significantly and the ratio between CEP and ML publications per year is far less pronounced (average ratio of 1 per 6).

[1] The focus is set on the SCOPUS database since it references a wider range of publishers (contrary to ScienceDirect or IEEE Xplore), including the ISCRAM digital library.
[2] Data files are available upon request.

The 10 duplicate results between R1 and R3, R1 and R5, and R3 and R5 queries (focus on the CEP approach) are the same 10 documents resulting from the R5 query (CEP × DSS/EWS × crisis management). Half of these 10 documents are authored by the ISCRAM community (in ISCRAM conferences, journals, etc.). They represent 32% of the R1 and 8.7% of the R3 results. The 24 duplicates found between R2 and R4, R2 and R6, and R4 and R6 (focus on the ML approach) are the same documents as the 24 results of the R6 query (ML × DSS/EWS × crisis management). They represent almost 7.3% of the R2 and 0.34% of the R4 results. A few of these 24 publications were authored by the ISCRAM community, not necessarily within the ISCRAM conferences series. R3 and R4 share 7 results, while there is no overlap in R1 with R2 and R5 with R6.

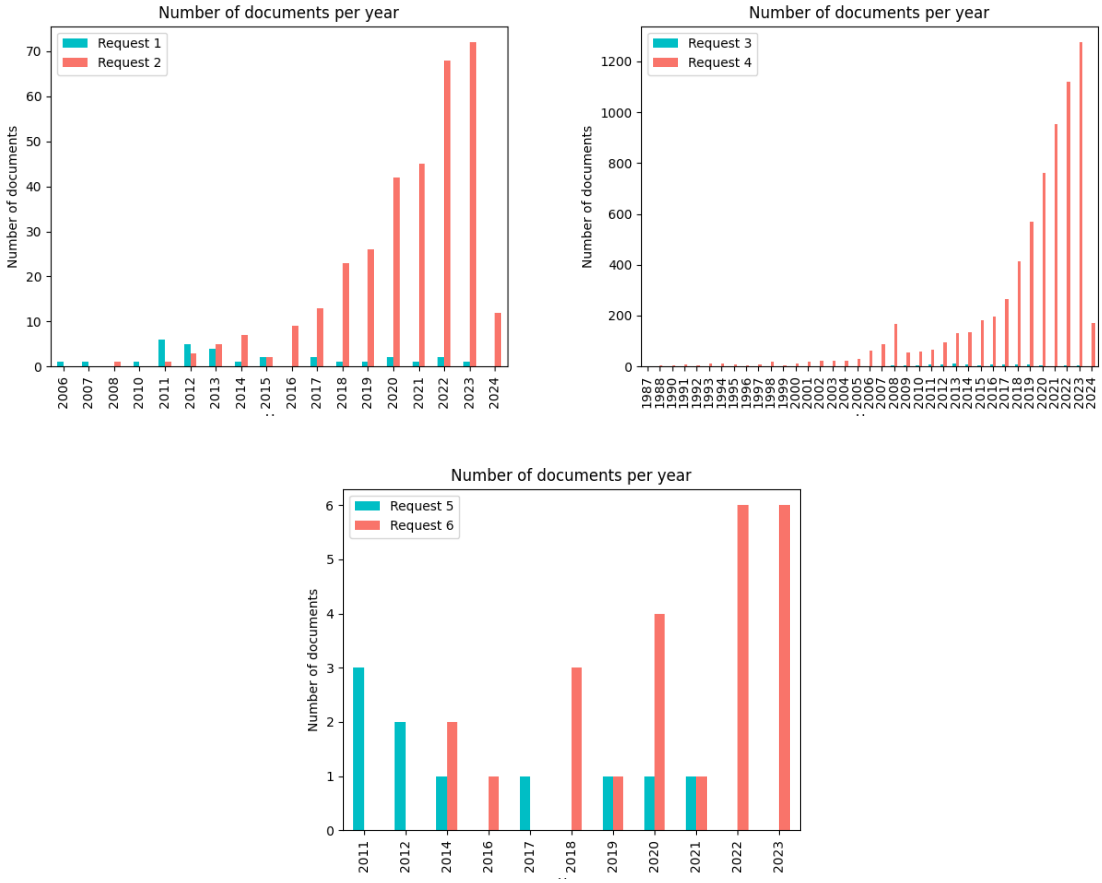


Figure 3. Distribution of the publications per year (R1 to R6 raw results)

Table 4. Duplicates between two queries raw results

Req.	R1	R2	R3	R4	R5
R2	0	-	-	-	-
R3	10	-	-	-	-
R4	-	24	7	-	-
R5	10	-	10	-	-
R6	-	24	-	24	0

The main outcomes of this quick quantitative exploratory data analysis are the following:

- ML is exponentially trending in DSS and EWS studies, exponentially overshadowing CEP for a decade.
- Within the crisis management application domain, the global ratio between CEP and ML publications is of 1 per 9 (R1 vs R2) up to 1 per 2.5 (R5 vs R6).

Deep Dive into the Research Topics

The previous findings show us that it will be more efficient to explore the topics of the publications retrieved by the R5 query and the R6 query to understand the use of CEP or ML for EWS and DSS in crisis management, as well as the seven results common to the R3 and R4 queries to understand the simultaneous presence of CEP and ML in these papers.

This section deep dives into the content of the retrieved documents. First, the relevance is checked by reading the title, abstract, introduction, and conclusion: publications outside the scope of this paper (data analytics to ultimately feed an EWS or a DSS) are not included in the analysis. Then, they are fully read to analyse the research proposal and results and, if applicable, the TRL (Technology Readiness Level) and the assumptions. These will help to understand the current trends found in the previous section, but also the gaps and the maturity level of the proposed solutions.

The results of the R5 query (Barthe-Delanoë et al., 2014; Braun et al., 2012; Fertier et al., 2020, 2021; Itria et al., 2017; Kovordanyi et al., 2012; Kraus et al., 2011; Mijović et al., 2011, 2019; Vranes et al., 2011) all show that the focus of these research works was on proving and demonstrating the integration of CEP components with other information system components (DSS, EWS, simulation tools, COP) within a crisis management context. Several information system architectures based on a hybrid ED-SOA (Event-Driven Service-Oriented Architecture) are also proposed to ensure the low coupling of these components. The proofs-of-concept and demonstrators presented in these publications are TRL 4 up to 6 by the end of the 2010s.

Out of the 24 results of query R6, 10 publications are excluded due to language constraints (despite automated translation efforts), being out of the scope of the study (e.g. wildlife activities) or having too few details about the processed data and/or the end use. The remaining 14 publications can be classified as follows, based on their authors' envisioned use:

- ML for DSS: emergency care triage (Rostam Niakan Kalhori, 2022), building structural assessment (Rahman et al., 2020), resources reallocation (Khouj et al., 2014, 2018), nuclear emergency management (Ionescu & Scheuermann, 2016), earthquake damage assessment (Ahadzadeh & Malek, 2021), contingency plans design (Kunikowski et al., 2020).
- ML for EWS: financial crisis (Zhao et al., 2023), citizen complaints (Zhang et al., 2023), flood forecast (Burrichter et al., 2023), building structural assessment (Kim et al., 2022).

These works all have a low TRL (evaluated at 3): the limitations and discussion sections show that the proofs-of-concept are still in their early stages. This can be explained by the recent technological hardware advances that made the wide use of ML algorithms possible. However, they all face the data input issue, either regarding the available volume, the available quality, or the available diversity of the data sources. Most of the authors underline the need for training, testing, and validating on (i) real data rather than on generated data and (ii) various datasets to overcome the overfitting bias.

Among the seven publications in common between the R3 query and R4 query results, one was recently retracted, and only two use both CEP and ML approaches. The works of (Bertoncelli et al., 2023; Sadanandan et al., 2022) aim to detect patterns in real-time data, using pre-defined CEP rules and historical data to learn data patterns with ML algorithms. Both CEP and ML are used concurrently rather than by hybridisation.

PAST AND CURRENT ISSUES

Integrating CEP in Information Systems Is A Mature Research Topic

It is interesting to note that despite a clear difference in the goal of each approach towards data processing and analytics, research on CEP has been almost abandoned in the information systems for crisis management domain since the late 2010s.

Given the findings in the previous section, it seems that this is mainly due to the maturity level of the research results (Barthe-Delanoë et al., 2014; Braun et al., 2012; Fertier et al., 2020, 2021; Itria et al., 2017; Kovordanyi et al., 2012; Kraus et al., 2011; Mijović et al., 2011, 2019; Vranes et al., 2011). The common topic of these papers was proving the benefits of integrating a CEP engine to feed an EWS or a DSS in a crisis management context with quite a high TRL. It means that they were ready for the next stage of development, which is production rather than research and proof.

However, while CEP has now been adopted as a powerful approach to analyze large volume of streaming data in real-time while operating with limited computational resources, a strong assumption was made in the publications

related to information systems for crisis management. It was assumed that the knowledge necessary to create and implement the CEP rules was available so that it was ready to use. The questions of sharing explicit knowledge, making tacit knowledge explicit, the availability of experts for knowledge gathering and CEP rules assessment, the volume of knowledge to deal with, etc., to create these CEP rules have not yet been addressed by the ISCRAM community.

Automatic Creation of CEP Rules

The SCOPUS query R7 (Table 5) looks for publications about the automatic generation of CEP rules without limit of time or application domain: 48 documents were retrieved, authored by the Computer Science community, who seized this question (Alakari et al., 2020; Margara et al., 2014; Mehdiyev et al., 2015) mainly in the context of physical/cyber intrusion detection, air quality monitoring, traffic monitoring.

Table 5. Set of SCOPUS queries and results

Req.	Content	Results
R7	TITLE-ABS-KEY("event process*" AND autom* W/5 rule*)	48 documents
R8	TITLE-ABS-KEY "event process*" AND "of thing*" AND autom* W/5 rule*)	14 documents

Among these 48 documents, 25% are related to the Internet of Things (IoT) and Web of Things (WoT) research areas (query R8). The distribution of the publications over time (Figure 4) shows that the question of the automatic creation of CEP rules is still an open issue and has been seized by the IoT and WoT communities in the last decade. This question has been on the roadmap of these research communities (Mehdiyev et al., 2015), among the first to adopt CEP in their software architectures.

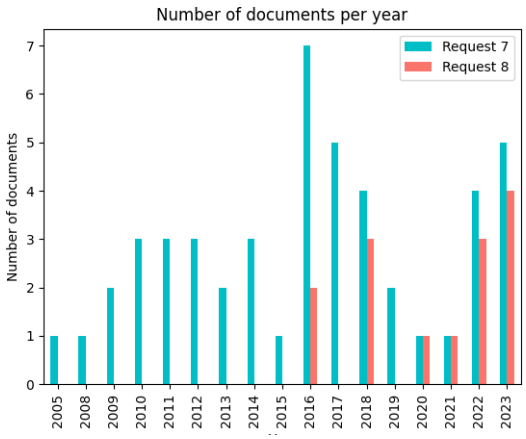


Figure 4. Distribution of the publications per year (R7 and R8)

Most of the papers propose using ML algorithms to generate CEP rules automatically: Deep Learning algorithms (Liu et al., 2022; Simsek et al., 2021; Simsek & Özdemir, 2018), classification algorithms (Kumar et al., 2023; Petersen et al., 2018; Sun et al., 2020), Principal Component Analysis (PCA) (Roldán-Gómez et al., 2023), time series shapelet (Li et al., 2019; Mousheimish et al., 2017). Only a few papers use approaches such as optimisation: evolutionary algorithms (Lv et al., 2022) or swarm algorithms (Bruns et al., 2023). All these research works focus on numerical and categorical data, mostly in a supervised training context.

CEP and ML are, in fact, complementary, given their features (see Table 1). CEP natively offers the processing of event flows in real time and enables patterns to be detected in real time, with limited computational resources. ML, on the other hand, makes it possible to exploit historical data to generate the CEP rules automatically (or as much as possible).

However, the results of the IoT and WoT communities may not be directly transposable to the field of crisis management (civil, humanitarian, agricultural, etc.). To prepare or to respond to a crisis, data are collected from various sources and in various types: numerical measures from sensors, text and images from social media (Coche et al., 2020; Power et al., 2022; Salemi et al., 2023) and from reports and press articles (Chasseray et al., 2021), speech (Cámbara et al., 2023), etc.

In other words, there is a significant challenge in processing structured and unstructured data when the CEP-ML solutions referenced above use datasets with numerical features (mostly) and categorical features.

Another key challenge is the learning part itself: most of the cited works use supervised algorithms, meaning their training datasets are labelled. In the context of crisis management, where the context is dynamic and the data sources volatile, the task of labelling datasets to train ML algorithms is unrealistic. It is then necessary to explore ways to perform CEP and ML hybridisation that rely on unsupervised learning algorithms.

CONCLUSION AND RECOMMENDATIONS

In this position paper, a short quantitative analysis was conducted on the papers related to CEP and ML for processing and analysing data to feed DSS and EWS. The naïve thought of CEP being overshadowed by the trending ML in the ISCRAM community is due to the maturity of the past research topics. However, the major issue of the automatic CEP rule creation was not addressed in the crisis management area. It is still an open issue, now seized by the Internet and Web of Things communities. They propose to use ML to deduce CEP rules, mainly using labelled datasets.

It is then recommended that the information systems for the crisis management community (particularly the ISCRAM community) focus on the potential of CEP in conjunction with ML. A hybrid approach may be the key to overcoming the challenges associated with automatic rule generation by exploiting the advantages of ML to process historical data and extract the necessary knowledge, whether explicit or tacit, whether from experts' interviews or written documents, images, etc.

It is crucial to address the assumptions related to the availability of knowledge and the resources required in the proposed solutions and to keep in mind the specific challenges faced by the crisis management domain compared to other application domains: the variety and the reliability of data sources; the availability of data sources and knowledge resources; structured, semi-structured and unstructured data; labelled and unlabeled data (to name a few).

A potential agenda for future research could lie in these four threads:

- Exploring methods to generate CEP rules from unsupervised ML algorithms automatically, trained on non-structured data;
- Optimizing CEP rules to balance accuracy, completeness, and efficiency considering the complexity and volatility of data sources;
- Adjusting CEP rules dynamically based on feedback from practitioners and end-users;
- Considering the ethical and societal implications of using ML-generated CEP rules such as privacy, transparency, bias, and fairness awareness (to name a few).

By investing in CEP-ML hybrid approaches adapted to the crisis management domain characteristics, DSS and EWS could significantly improve responsiveness and accuracy in crisis situations.

REFERENCES

Ahadzadeh, S., & Malek, M. R. (2021). Earthquake damage assessment based on user generated data in social networks. *Sustainability (Switzerland)*, *13*(9). Scopus. https://doi.org/10.3390/su13094814

Alakari, A., Li, K. F., & Gebali, F. (2020). A situation refinement model for complex event processing. *Knowledge-Based Systems*, 105881. https://doi.org/10.1016/j.knosys.2020.105881

Barthe-Delanoë, A.-M., Truptil, S., & Bénaben, F. (2014). Agility of crisis response: Gathering and analyzing data through an event-driven platform. *Proceedings of the 11th International ISCRAM Conference*. 11th International Conference on Information Systems for Crisis Response and Management, University Park, Pennsylvania, USA.

Bertoncelli, C. M., Costantini, S., Persia, F., Bertoncelli, D., & D'Auria, D. (2023). PredictMed-epilepsy: A multi-agent based system for epilepsy detection and prediction in neuropediatrics. *Computer Methods and Programs in Biomedicine*, *236*. Scopus. https://doi.org/10.1016/j.cmpb.2023.107548

Braun, A., Kroner, P., Leontyeva, Y., & Siller, D. (2012). Event processing in sensor networks: A solution for integrated emergency management. *Proceedings of the 6th ACM International Conference on Distributed Event-Based Systems*, 371–372. https://doi.org/10.1145/2335484.2335527

Bruns, R., Dunkel, J., & Seremet, S. (2023). Learning Ship Activity Patterns in Maritime Data Streams: Enhancing CEP Rule Learning by Temporal and Spatial Relations and Domain-Specific Functions. *IEEE Transactions on Intelligent Transportation Systems*, *24*(10), 11384–11395. https://doi.org/10.1109/TITS.2023.3282246

Burrichter, B., Hofmann, J., Koltermann da Silva, J., Niemann, A., & Quirmbach, M. (2023). A Spatiotemporal Deep Learning Approach for Urban Pluvial Flood Forecasting with Multi-Source Data. *Water*, *15*(9), Article 9. https://doi.org/10.3390/w15091760

Cámbara, G., Grivolla, J., Farrús, M., & Wanner, L. (2023). *Automatic Speech Translation for Multinational First Responder Teams*. *2023-text*, 188–196. Scopus.

Chasseray, Y., Barthe-Delanoë, A.-M., Negny, S., & Le Lann, J.-M. (2021). Automated unsupervised ontology population system applied to crisis management domain. *Proc. Int. ISCRAM Conf.*, *2021-May*, 968–981. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85121216060&partnerID=40&md5=b859541d5988c619bbb9bc50c26785f5

Coche, J., Montarnal, A., Tapia, A., & Benaben, F. (2020). Automatic Information Retrieval from Tweets: A Semantic Clustering Approach. *ISCRAM 2020 - 17th International Conference on Information Systems for Crisis Response and Management*, p.134-141. https://hal.telecom-paris.fr/hal-02926851

Domala, J., Dogra, M., Masrani, V., Fernandes, D., D'souza, K., Fernandes, D., & Carvalho, T. (2020). Automated Identification of Disaster News for Crisis Management using Machine Learning and Natural Language Processing. *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 503–508. https://doi.org/10.1109/ICESC48915.2020.9156031

Etzion, O., & Niblett, P. (2010). *Event Processing in Action* (1st ed.). Manning Publications Co.

Fertier, A., Barthe-Delanoë, A.-M., Montarnal, A., Truptil, S., & Bénaben, F. (2020). A new emergency decision support system: The automatic interpretation and contextualisation of events to model a crisis situation in real-time. *Decision Support Systems*, 113260. https://doi.org/10.1016/j.dss.2020.113260

Fertier, A., Martin, G., Barthe-Delanoë, A.-M., Lesbegueries, J., Montarnal, A., Truptil, S., Bénaben, F., & Salatgé, N. (2021). Managing events to improve situation awareness and resilience in a supply chain. *Computers in Industry*, *132*. https://doi.org/10.1016/j.compind.2021.103488

Fülöp, L. J., Beszédes, Á., Tóth, G., Demeter, H., Vidács, L., & Farkas, L. (2012). Predictive complex event processing: A conceptual framework for combining complex event processing and predictive analytics. *Proceedings of the Fifth Balkan Conference in Informatics*, 26–31. https://doi.org/10.1145/2371316.2371323

Ionescu, T. B., & Scheuermann, W. (2016). Improving the reliability of decision-support systems for nuclear emergency management by leveraging software design diversity. *Journal of Computing and Information Technology*, *24*(1), 45–63. Scopus. https://doi.org/10.20532/cit.2016.1002700

Itria, M. L., Kocsis-Magyar, M., Ceccarelli, A., Lollini, P., Giunta, G., & Bondavalli, A. (2017). Identification of critical situations via Event Processing and Event Trust Analysis. *Knowledge and Information Systems*, *52*(1), 147–178. https://doi.org/10.1007/s10115-016-1009-x

Jaquart, P., Dann, D., & Weinhardt, C. (2021). Short-term bitcoin market prediction via machine learning. *The Journal of Finance and Data Science*, *7*, 45–66. https://doi.org/10.1016/j.jfds.2021.03.001

Khouj, M. T., Alsubaie, A., Alutaibi, K., Ahmed, H. M., Sarkaria, S., & Martí, J. R. (2018). Intelligent decision system for responsive crisis management. *International Journal of Critical Infrastructures*, *14*(4), 375–399. Scopus. https://doi.org/10.1504/IJCIS.2018.095619

Khouj, M. T., Sarkaria, S., & Marti, J. R. (2014). Decision assistance agent in real-time simulation. *International Journal of Critical Infrastructures*, *10*(2), 151–173. Scopus. https://doi.org/10.1504/IJCIS.2014.062968

Kim, J., Hatzis, J. J., Klockow, K., & Campbell, P. A. (2022). Building Classification Using Random Forest to Develop a Geodatabase for Probabilistic Hazard Information. *Natural Hazards Review*, *23*(3). Scopus. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000561

Kovordanyi, R., Pelfrene, J., & Eriksson, H. (2014). Supporting Instructors' Decision Making in Simulator-Based Training for Crisis Management. *Proceedings of the 11th International ISCRAM Conference*. 11th International Conference on Information Systems for Crisis Response and Management, University Park, Pennsylvania, USA.

Kovordanyi, R., Pelfrene, J., Rankin, A., Schreiner, R., Jenvald, J., Morin, M., & Eriksson, H. (2012). *Real-time Support for Exercise Managers' Situation Assessment and Decision Making*. http://liu.diva-portal.org/smash/record.jsf?pid=diva2:546538

Kraus, L., Stanojević, M., Tomašević, N., & Mijović, V. (2011). A decision support system for building evacuation based on the EMILI SITE environment. *Proceedings of the 2011 20th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2011*, 334–336. https://doi.org/10.1109/WETICE.2011.59

Kumar, S. S., Kumar, A., Chandra, R., Agarwal, S., Syafrullah, M., & Adiyarta, K. (2023). Rule Extraction using Machine Learning Classifiers for Complex Event Processing. *2023 10th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 355–360. https://doi.org/10.1109/EECSI59885.2023.10295840

Kunikowski, G., Kosieradzka, A., & Kąkol, U. (2020). Methodology for the preparation of rescue plans – a Polish case study. *Disaster Prevention and Management: An International Journal*, *29*(3), 313–321. Scopus. https://doi.org/10.1108/DPM-02-2019-0062

Li, Y., Zhang, T., & Song, C. (2019). Research on Generation Algorithm of Complex Event Processing Rules Based on Time Series. *2019 Eleventh International Conference on Advanced Computational Intelligence (ICACI)*, 124–128. https://doi.org/10.1109/ICACI.2019.8778586

Lima, M., Lima, R., Lins, F., & Bonfim, M. (2022). Beholder – A CEP-based intrusion detection and prevention systems for IoT environments. *Computers & Security*, *120*, 102824. https://doi.org/10.1016/j.cose.2022.102824

Liu, Y., Yu, W., Gao, C., & Chen, M. (2022). An Auto-Extraction Framework for CEP Rules Based on the Two-Layer LSTM Attention Mechanism: A Case Study on City Air Pollution Forecasting. *Energies*, *15*(16), Article 16. https://doi.org/10.3390/en15165892

Luckham, D. C. (2011). *Event Processing for Business: Organizing the Real-Time Enterprise*. John Wiley & Sons Inc.

Lv, J., Yu, B., & Sun, H. (2022). CEP Rule Extraction Framework Based on Evolutionary Algorithm. *2022 11th International Conference of Information and Communication Technology (ICTech))*, 245–249. https://doi.org/10.1109/ICTech55460.2022.00056

Margara, A., Cugola, G., & Tamburrelli, G. (2014). Learning from the Past: Automated Rule Generation for Complex Event Processing. *Proceedings of the 8th ACM International Conference on Distributed Event-Based Systems*, 47–58. https://doi.org/10.1145/2611286.2611289

Mehdiyev, N., Krumeich, J., Enke, D., Werth, D., & Loos, P. (2015). Determination of Rule Patterns in Complex Event Processing Using Machine Learning Techniques. *Procedia Computer Science*, *61*, 395–401. https://doi.org/10.1016/j.procs.2015.09.168

Mijović, V., Tomašević, N., Janev, V., Stanojević, M., & Vraneš, S. (2011). Ontology enabled decision support system for emergency management at airports. *Proceedings of the 7th International Conference on Semantic Systems*, 163–166. https://doi.org/10.1145/2063518.2063540

Mijović, V., Tomašević, N., Janev, V., Stanojević, M., & Vraneš, S. (2019). Emergency Management in Critical Infrastructures: A Complex-Event-Processing Paradigm. *Journal of Systems Science and Systems Engineering*, *28*(1), 37–62. https://doi.org/10.1007/s11518-018-5393-5

Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.

Mousheimish, R., Taher, Y., & Zeitouni, K. (2017). Automatic Learning of Predictive CEP Rules: Bridging the Gap between Data Mining and Complex Event Processing. *Proceedings of the 11th ACM International Conference on Distributed and Event-Based Systems*, 158–169. https://doi.org/10.1145/3093742.3093917

Petersen, E., Antonio To, M., Maag, S., & Yamga, T. (2018). An Unsupervised Rule Generation Approach for Online Complex Event Processing. *2018 IEEE 17th International Symposium on Network Computing and Applications (NCA)*, 1–8. https://doi.org/10.1109/NCA.2018.8548210

Pottebaum, J., Artikis, A., Marterer, R., Paliouras, G., & Koch, R. (2011). Event definition for the application of event processing to intelligent resource management. *International Conference on Information Systems for Crisis Response and Management (ISCRAM)*. http://users.iit.demokritos.gr/~paliourg/papers/ISCRAM2011.pdf

Power, R., Robinson, B., & Cameron, M. (2022). *Insights from a Decade of Twitter Monitoring for Emergency Management*. *2022-November*, 247–257. Scopus.

Rahman, T., Ahmed, Y., Alam, T., Shakil, M. H. R., Hossain, M. T., Khan, F. N., Dipto, S. M., Akhond, M. R., & Alam, M. A. (2020). *Evaluation of Earthquake Resistance of Urban Buildings using Image Processing and Machine Learning Techniques*. 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2020. Scopus. https://doi.org/10.1109/CSDE50874.2020.9411582

Roldán-Gómez, J., Boubeta-Puig, J., Carrillo-Mondéjar, J., Castelo Gómez, J. M., & del Rincón, J. M. (2023). An automatic complex event processing rules generation system for the recognition of real-time IoT attack patterns. *Engineering Applications of Artificial Intelligence*, *123*, 106344. https://doi.org/10.1016/j.engappai.2023.106344

Rostam Niakan Kalhori, S. (2022). Towards the Application of Machine Learning in Emergency Informatics. *Studies in Health Technology and Informatics*, *291*, 3–16. Scopus. https://doi.org/10.3233/SHTI220003

Sadanandan, S. K., Ahmed, A., Pandey, S., & Srivastava, A. K. (2022). Synchrophasor data analytics for anomaly and event detection, classification, and localization. In *Intelligent Data Mining and Analysis in Power and Energy Systems: Models and Applications for Smarter Efficient Power Systems* (pp. 105–127). Scopus. https://doi.org/10.1002/9781119834052.ch5

Salemi, H., Senarath, Y., & Purohit, H. (2023). *A Comparative Study of Pre-trained Language Models to Filter Informative Code-mixed Data on Social Media during Disasters*. *2023-text*, 920–932. Scopus.

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://doi.org/10.1147/rd.33.0210

Shah, N., Shah, S., Jain, P., & Doshi, N. (2022). Overview of Present-Day IoT Data Processing Technologies. *Procedia Computer Science*, *210*, 277–282. https://doi.org/10.1016/j.procs.2022.10.150

Simsek, M. U., & Özdemir, S. (2018). CEP Rule Extraction From Unlabeled Data in IoT. *2018 3rd International Conference on Computer Science and Engineering (UBMK)*, 429–433. https://doi.org/10.1109/UBMK.2018.8566255

Simsek, M. U., Yildirim Okay, F., & Ozdemir, S. (2021). A deep learning-based CEP rule extraction framework for IoT data. *The Journal of Supercomputing*, *77*(8), 8563–8592. https://doi.org/10.1007/s11227-020-03603-5

Sun, Y., Li, G., & Ning, B. (2020). Automatic Rule Updating based on Machine Learning in Complex Event Processing. *2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS)*, 1338–1343. https://doi.org/10.1109/ICDCS47774.2020.00176

UNDRR. (2007). *Early warning system | Sendai Framework Terminology on Disaster Risk Reduction* [Institutional]. United Nations Office for Disaster Risk Reduction. http://www.undrr.org/terminology/early-warning-system

Vranes, S., Stanojevic, M., Janev, V., Mijovic, V., Tomasevic, N., Kraus, L., & Ilic, Z. (2011). Application of Complex Event Processing Paradigm in Situation Awareness and Management. *2011 22nd International Workshop on Database and Expert Systems Applications*, 289–293. https://doi.org/10.1109/DEXA.2011.15

Zhang, Z., Lin, X., & Shan, S. (2023). Big data-assisted urban governance: An intelligent real-time monitoring and early warning system for public opinion in government hotline. *Future Generation Computer Systems*, *144*, 90–104. Scopus. https://doi.org/10.1016/j.future.2023.03.004

Zhao, Z., Li, D., & Dai, W. (2023). Machine-learning-enabled intelligence computing for crisis management in small and medium-sized enterprises (SMEs). *Technological Forecasting and Social Change*, *191*, 122492. https://doi.org/10.1016/j.techfore.2023.122492

PROJECT SYNERGISE - increased safety and efficiency of first responder operations

ABSTRACT

SYNERGISE is an EU-funded research project aimed at enhancing the safety and efficiency of first responder operations through innovative technological solutions. With the evolving landscape of emergencies and disasters, the need for effective response strategies has become increasingly critical. SYNERGISE addresses this challenge by leveraging advanced technologies to empower first responders with enhanced capabilities and tools.

The project focuses on developing integrated robotic solutions and intelligent decision-support systems tailored to the needs of first responders. By harnessing the power of robotics, artificial intelligence, human-machine collaboration and data analytics, SYNERGISE aims to revolutionize the way emergency situations are managed and responded to.
Key objectives of the project include:

1. Designing and implementing robotic platforms capable of assisting first responders in various tasks, such as search and rescue, reconnaissance, and hazardous material/environment handling.
2. Developing intelligent decision-support systems that analyze real-time data from multiple sources to provide actionable insights and recommendations to first responders.
3. Enhancing collaboration and coordination among different response teams through advanced communication and information-sharing technologies.
4. Conducting rigorous testing and validation of the developed solutions in real-world scenarios to ensure their effectiveness and reliability.

Through collaborative efforts involving researchers, industry partners, and emergency response organizations, SYNERGISE aims to deliver impactful solutions that enhance the safety of first responders and improve the efficiency of their deployments. By embracing innovation and technology, SYNERGISE seeks to pave the way for a safer and more resilient future in emergency response operations.

PRESENTER

Tiina Ristmäe
Tiina.ristmaee@thw.de
German Government Agency for Technical Relief

Vangelis Sdongos
e.sdongos@astrial.de
ASTRIAL GmbH

DIGITAL EMERGENCY COMMUNICATION IN THE NORDIC COUNTRIES

ABSTRACT

During disasters and emergencies, reliable means of communication are of vital importance to life- saving efforts. This poster presents findings from DIGeMERGE, which is a research project on Digital Emergency Communication in four highly digitalized countries: Norway, Sweden, Finland, and Denmark. DIGeMERGE is funded by the CHANSE (Collaboration of Humanities and Social Sciences in Europe) call for projects on Transformations: Social and cultural dynamics in the digital age. The project consortium consists of the Peace Research Institute Oslo (PRIO), Malmö University in Sweden, the IT University of Copenhagen in Denmark, and the Centre for Machine Vision and Signal Analysis, University of Oulu in Finland.

The overarching objective of the DIGeMERGE project is to understand both the everyday practical challenges of emergency management digitalization and the deeper implications of these developments for the state-citizen relationship. The project pursues this objective by investigating challenges associated with new digital emergency communication tools and related debates about societal and individual responsibility, protection needs, vulnerability and rights in the four Nordic countries.

The governments of the Nordic countries highlight the need for information sharing, cooperative initiatives and technical innovation to improve emergency response capabilities and readiness, focusing on the adoption of cell broadcast alert systems and mobile broadband networks (3GPP, LTE, and 5G). There is also a strong attention to the need to set standards that provide smooth communication in cross-border emergencies and ensures interoperability across countries.

In the poster for ISCRAM, we will present a case study of the implementation of the cell broadcast emergency alert system in Norway, which is a new system designed as a digital supplement to the existing post-World War II air-raid typhoons or sirens. At least 2.2 million of Norway's 5.5 million people live too far away from the nearest siren to be able to hear it. In the latest test of the cell broadcast alert for mobile phones ("Nødvarsel på mobil"), approximately 90% of the population received the alert on their phone. Unlike the typhoons or sirens, which are operated by Norway's Civil Defence and were designed primarily for war-time emergencies, the cell broadcast alert system is operated by the police in peace-time and the Civil Defence in war-time. The Norwegian Police Directorate has developed a set of templates that consist of potential scenarios where they foresee the alert system to be put into use. The poster will present examples of these templates to show how the new technology is transforming the way authorities communicate with the public during emergencies, and will also give examples of reactions on social media to the use of the system in a real-life emergency - the extreme weather event "Hans" - illustrating the potential impact of this technological and organizational transformation on expectations about emergency management in Norway.

For more information about the DIGeMERGE project, see: https://www.prio.org/projects/DIGeMERGE

PRESENTER/DEMONSTRATOR

Åshild Kolås* ashild@prio.org
Peace Research Institute Oslo (PRIO)

Fuzel Shaik
Fuzel.Shaik@student.oulu.fi
University of Oulu

Development of a Responding Laboratory for the Collection of Data on Major Emergency Incidents

ABSTRACT

As part of the incident management process, it is necessary to collect data on the event, its course and how it is managed. This data can be used after the incident to continuously improve the entire incident management process. At the same time, it is necessary to be able to collect valid as well as precise data quickly during the incident and to present it appropriately for decision support to initiate measures in order to deal with the incident. The poster shows the development and the use of a laboratory that serves these purposes. The RespondingLab enables researchers to collect data during exercises or in real-life events, allowing them to evaluate processes and structures. The RespondingLab also serves as a kind of test bed for new technologies. For example, different sensors on drones and various options for visualizing the situation are currently being tested. To date, the laboratory has been used, for example, for the construction of a temporary bridge after a flood event. In this case, the laboratory provided 3D data of the terrain as well as live data on the bridge construction.

The RespondingLab is supplemented by a vehicle and a trailer so that the laboratory's technologies can be used flexible in different locations. The lab's sensors can be set up quickly, and they are just as quickly connected to each other via the lab's own ad hoc WLAN. The poster shows which sensors and technologies the laboratory uses to collect data and how this data can be analyzed. The poster is intended to give researchers an idea of how data from emergency incidents can be collected and analyzed in the field. This research method not only supports the improvement of processes and structures for first responders, but also the active provision of analyzed data during real emergency incidents.

PRESENTER/DEMONSTRATOR

Lennart Landsberg
lennart_walter.landsberg@th-koeln.de
TH Köln – University of Applied Sciences / Institute of Rescue Engineering and Civil Protection

Thomas Säger (Conference participation)
thomas.saeger@th-koeln.de
TH Köln – University of Applied Sciences / Institute of Rescue Engineering and Civil Protection

Konrad Barth
konrad.barth@th-koeln.de
TH Köln – University of Applied Sciences / Institute of Rescue Engineering and Civil Protection

Ompe Aimé Mudimu

ompe_aime.mudimu@th-koeln.de
TH Köln – University of Applied Sciences / Institute of Rescue Engineering and Civil Protection

Enhancing Emergency Preparedness in Mass Gatherings

ABSTRACT

This study emerges from a vital collaboration with the firefighters in France (i.e. Service Départemental d'Incendie et de Secours, SDIS), underscoring the critical role of emergency services in fighting fires, providing emergency assistance, and preventing accidents and disasters, particularly in the context of mass gatherings. The primary purpose of our research is to enhance preparedness and proactive response planning for medical emergencies during large-scale events through a sophisticated crowd simulation and analysis approach.

Our approach unfolds in two layers. The first layer, **modeling and assessment**, focuses on creating models of different urgency levels of projected scenarios, including their cascading effects. It aims to assess the agility and effectiveness of various emergency response strategies and their capability to mitigate such cascading phenomena. The second layer, **analysis, and decision-making**, is dedicated to identifying critical thresholds where medical responses might be overwhelmed. It seeks to optimize the allocation of medical resources and improve response times, ensuring agile and efficient emergency handling.

A practical application of our approach was demonstrated at a major music festival held on the former air base of Cambrai, France. We employed AnyLogic's pedestrian library and agent-based modeling to develop a crowd simulation that fulfills two expectations: visualizing and monitoring crowd flow within confined spaces and anticipating the impacts of current and future risks and their cascading effects on both the crowd and emergency responses.

Our outcomes to date reveal three significant capabilities of the simulation:

- Simulating normal crowd movement within a realistically scaled environment.
- Modeling low urgency emergency scenarios, allowing us to compute crucial metrics like crowd density and medical response time.
- Illustrating high urgency medical emergencies and their potential to incite secondary incidents, such as widespread panic, and examining how crowd movement patterns might create barriers or bottlenecks that impede emergency responses.

An ongoing effort is directed towards simulating how secondary incidents propagate through the crowd, affecting overall safety and the efficacy of response strategies. Regarding an outlook for future work, we aim to refine our simulation capabilities to encompass a broader spectrum of emergency scenarios and their complex interdependencies. This includes enhancing the model's accuracy in predicting cascading effects and integrating real-time data for dynamic scenario planning.

PRESENTER/DEMONSTRATOR

Shortly state the previous experience of the presenter(s) at ISCRAM or related meetings, and any other qualifications for this responsibility, such as recent papers on the topic you propose or experience in using the tools or procedures you want to demonstrate. In addition, you should note which presenter plans to be at the conference.

Jiayao L
jiayao.li@mines-albi.fr
IMT Mines Albi – Carmaux

Aurélie Congès
aurelie.conges@mines-albi.fr
IMT Mines Albi – Carmaux

Frédérick Bénaben>
frederick.benaben@mines-albi.f
IMT Mines Albi – Carmaux

Community-Based Smart City Digital Twin Platform for Optimised DRM operations and Enhanced Community Disaster Resilience

ABSTRACT

PANTHEON is a project funded by the European Union's Horizon Europe programme, bringing together 17 partners from 9 different countries under the research theme Disaster-Resilient Society for Europe. The project started on 1 January 2023 and its planned durations is 36 months. The PANTHEON consortium aims to develop a community-based Smart City Digital Twin that facilitates disaster risk management operations while reducing vulnerability and strengthening community resilience.

The implementation of technologies is impossible without considering the context in which they will be deployed, thus the first year of the project has focused on the study of the PANTHEON ecosystem. As a first step, a risk mapping and assessment was carried out to identify the most relevant disasters in the two pilot cities: Athens and Vienna. This assessment includes both natural and man-made hazards. In addition, the regulatory framework was analysed in depth, including national civil protection plans and disaster management strategies for the two countries, as well as EU-wide directives and initiatives. On the social side, an assessment of the vulnerability and capacity of the community in the event of disasters has been carried out, with a particular focus on vulnerable groups such as the elderly, children or people with physical or mental disabilities. One of the outputs was the development of a set of indicators to assess community capacity and vulnerability. In addition, a series of recommendations were produced to involve the community, especially vulnerable groups, in decision-making on disaster management issues. From the technological viewpoint, the technical and scientific capabilities that exist to support disaster planning and management have been studied, something that has greatly assisted in the creation of a roadmap of technologies that PANTHEON intends to develop as well as their expected improvements.

These results, together with a participatory design process, have been the basis for conceptualising the idea of the PANTHEON system. Thus, taking advantage of the diversity of data sources available in a smart city, such as satellites, sensors, infrastructure data, traffic data, UAV data and community data, PANTHEON aims to be a system composed of two fundamental elements, both focused on the pre-disaster phase. On one hand, PANTHEON will provide a training platform to simulate disaster situations and train for complex and economically difficult scenarios whereas, on the other, a planning and early warning platform will be integrated to allow end-users optimise and improve resource management and disaster planning, with a special focus on cascading effects.

To test the usability and functionality of the PANTHEON system, four scenarios have been created that reflect the most important disasters in the two cities, as well as the most relevant tools for each scenario. The training platform will be tested by visualising and simulating an earthquake near Athens and a fire caused by a cyber-attack on a power plant in Vienna. The planning and early warning platform will be tested by visualising and simulating a forest fire near Athens and the propagation of a heat wave in Vienna.

PRESENTER/DEMONSTRATOR

Dr. Lola Vallès Port*
lvalles@gencat.cat
Institute for Public Security of Catalonia

Kiril Shtefchyk Tatarchuk
Kiril.shtefchyk@gencat.cat
Institute for Public Security of Catalonia

Mike Karamousadakis
mike.karamousadakis@twi.gr TWI
HELLAS

John Tsaloukidis
j.tsaloukidis@kemea-research.gr
KE. ME. A

Automated Integration of Humanitarian Data into Route Planning to support Disaster Response

ABSTRACT

During disaster events, humanitarian organisations often need to plan routes to deliver goods and services to affected areas, identify evacuation paths, and to determine the reachability of services by the population. Especially in the case of flooding, roads can become blocked and impassible, meaning that vehicles need to be redirected from the usually optimal route. To efficiently plan suitable routes during disaster events, a routing service must consider the constantly evolving condition of the road network by integrating data from different sources and organisations.
Existing services such as Google Maps do not offer such capabilities.

To fill this gap, a prototype system has been developed called the HeiGIT Disaster Portal [2], which acts as a data storage and a proxy to services such as Openrouteservice for Disaster Management [1]. Data relating to disaster events (such as flooded regions) can be added to the portal, which is then stored centrally for access by other users and services. When requests to openrouteservice are made through the portal, these data can be automatically included in the request as areas to be avoided without the user needing to manually draw them.

The newest set of developments includes an advanced routing algorithm, that supports extensive and intricate avoidance zones, which would normally slow down calculations or render them impossible. The algorithm combines the centralized storage with rapid pre-processing steps for routing graphs, thereby ensuring efficient routing and reachability analysis on complex avoidance data.

As an example, the HeiGIT Disaster Portal was used during the Anatolia earthquake in 2023. Data relating to closed border crossings and damaged road infrastructure were provided to the portal by humanitarian organisations, which were then used to avoid areas in routing requests.

In the future, the Disaster Portal will be extended to include automatic ingestion of data from sources such as the Copernicus Emergency Management Service [3], as well as providing access to additional services and operations.

[1] "Openrouteservice for Disaster Management." HeiGIT gGmbH. Accessed May 19, 2023. https://disaster.openrouteservice.org/

[2] "HeiGIT Disaster Portal." HeiGIT gGmbH. Accessed May 19, 2023. https://disaster- portal.heigit.org/

[3] "Copernicus Emergency Management Service." Directorate Space, Security and Migration, European

Commission Joint Research Centre (EC JRC). Accessed March 14, 2024. https://emergency.copernicus.eu/

[1] "Openrouteservice for Disaster Management." HeiGIT gGmbH. Accessed March 14, 2024. https://disaster.openrouteservice.org/

[2] "HeiGIT Disaster Portal." HeiGIT gGmbH. Accessed March 14, 2024. https://disaster- portal.heigit.org/

[3] "Copernicus Emergency Management Service." Directorate Space, Security and Migration, European Commission Joint Research Centre (EC JRC). Accessed March 14, 2024. https://emergency.copernicus.eu/

PRESENTER

Julian Psotta*
julian.psotta@heigit.org
Heidelberg Institute for Geoinformation Technology (HeiGIT gGmbH) at the University of Heidelberg

Dr. Christina Ludwig
chrisina.ludwig@heigit.org HeiGIT gGmbH

Anne Schauss
anne.schauss@heigit.org HeiGIT
gGmbH

Decision Support Systems for Public Authorities to Prepare and Manage Food Shortages

ABSTRACT

Crises such as the COVID pandemic and Russia's war of aggression against Ukraine have shown that food shortages are more likely than previously thought. In such a case, the state must maintain the population's food supply to avoid a crisis. This implies that the relevant authorities should be able to intervene in the food supply system in a directive manner. To do this, they need knowledge of the structures and capabilities of the food industry and perhaps even up-to-date operational information such as stock levels, supply and demand estimates and forecasts.

The disruptions in the food supply chain during COVID and the Russian-Ukrainian war have shown that cooperation and active communication between the responsible authorities and the food industry are essential for managing and solving problems. Instruments for such systematic cooperation do not currently exist in Germany. The events of recent years have shown that the authorities' crisis management of food supply problems is currently limited to manual data queries and spontaneous ad-hoc measures.

For this reason, the aim of the project “KRITIS-ENV” is to develop suitable cooperation and decision support systems. This is intended to improve the authorities' ability to act in the event of potential food supply shortages to make the food system more resilient.

To this end, the project will develop three tools. First, the development of a Risk Management and Communication System (RMCS) creates a standardised platform to facilitate and simplify management and communication between authorities and food industry companies. Second, the development of a decision support system and a database of information on the food supply in the form of a Geographical Information System (GIS) improves the transparency of the food system for the authorities and thus increase their ability to act and react. And third, contingency plans for selected scenarios provide a system of procedures and responsibilities to support appropriate action.

Methodically, a literature review and interviews will be used to define the requirements for the RMCS. For the GIS system, a comprehensive dataset of the German food supply system will be established. This database will provide the basis for a vulnerability analysis of the German food supply system which will determine the design of a GIS-based decision support system for food shortages. The derivation of several contingency plans will be done by practitioners accompanied by researchers. This Action Based Research setting will also result in a process blueprint for plan development.

To test the developed tools, we will specify a scenario in which we assume the closure of a large number of points of sale (e.g. due to staff shortage). For this, we develop a multi-objective optimization problem to select the optimal subset of POS to maximize food accessibility and (at the same time) efficiency of logistics operations.

The project is carried out in close cooperation with the Lower Saxony Ministry of Food, Agriculture and Consumer Protection, the software company Calpana and Kühne Logistics University. The project is expected to start on the 1st of March 2024.

PRESENTER

Ina-Linda Deuchert *
Ina-Linda.Deuchert@klu.org
Kühne Logistics University

Prof. Dr. Hanno Friedrich
Hanno.Friedrich@klu.org Kühne
Logistics University

Ute Menski *
Ute.Menski@klu.org Kühne
Logistics University

TREEADS - Holistic Fire Management Ecosystem for Prevention, Detection, and Restoration of Environmental Disasters

ABSTRACT

Wildfires are a severe threat across Europe, causing significant environmental and economic damage. They are becoming more intense and widespread as a result of climate change, particular forestry practices, ecosystem deterioration, and rural depopulation. Longer fire seasons, more frequent fire events, an increase in fire-prone locations, and more extreme fire behaviour are all projected to enhance fire risk as a result of climate change. According to the European Forest Fire Information System (EFFIS), the wildfire season of 2022 was the second worst on record, in terms of burnt area (881 275 hectars) and number of fires [1]. The latest UNEP and GRID-Arendal report on wildfires foresees a global increase of extreme fire events by up to 14 percent by 2030 [2]. Extreme wildfire events, in addition to their devastating ecological impact, have an unparalleled social cost in terms of impacts on human health and life, as well as environmental and economic losses.

The TREEADS project addresses several major challenges that wildfires pose, including current technological and infrastructural restrictions and severe environmental, societal, and economic consequences [3]. The TREEADS fire management ecosystem will include various innovative technologies and systems covering all time-interconnected stages of wildfires (before, during, and after). Over 26 different technologies are optimized, developed, and validated in the implementation of TREEADS, including coordinated novel pilotless aircraft (multi-rotor and fixed wing) and satellite systems, an advanced analytical and modeling framework for fire prevention, detection, and response, forest and landscape restoration and an emerging set of fire protection and suppression technologies. Expected results are, among others, the minimization of the significant impacts caused by extreme fires and experience strengthening for end-users within a rich set of realistic scenarios via the project's training programme. In TREEADS, we focus our research, development, and integration efforts on delivering systems of high relevance and benefit to the relevant end-users and the forestry market while providing them, in turn, with an excellent opportunity to understand potential weaknesses or limitations of current technologies. The TREEADS tools and solutions are demonstrated and validated under actual operating conditions in eight complex pilot implementations in Taiwan and seven European countries: Austria, Norway, Spain, Italy, Romania, Greece, and Germany. The project is also carrying out a disciplined assessment of the impact of the developed technologies based on methods that consider both social and financial aspects (e.g., Social Return on Investment (S-ROI)).

Acknowledgment: TREEADS has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 101036926. The findings were published in several conferences such as RISE-SD2023, ICFFR-Coimbra and IAFSS2023.

References:

[1] J. San-Miguel-Ayanz, Advance Report on Forest Fires in Europe, Middle East and North Africa 2022, EUR 31479 EN, Publications Office of the European Union, Luxembourg, 2023, ISBN 978-92-68-02143-9, doi:10.2760/091540, JRC133215.

[2] United Nations Environment Programme (2022). Spreading like Wildfire – The Rising Threat of Extraordinary Landscape Fires. A UNEP Rapid Response Assessment. Nairobi. Available: https://www.unep.org/resources/report/spreading-wildfire-rising-threat-extraordinary-landscape-fires

[3] https://treeads-project.eu/

PRESENTER/DEMONSTRATOR

Kemal S. Arsava.*
kemal.sarp.arsava@risefr.no
RISE Fire Research, Trondheim, Norway

Ragni F. Mikalsen
ragni.mikalsen@risefr.no
RISE Fire Research, Trondheim, Norway

Ellen Synnøve Skilbred
ellen.synnove.skilbred@risefr.no
RISE Fire Research, Trondheim, Norway

Janne S. Fjærestad
janne.siren.fjarestad@risefr.no
RISE Fire Research, Trondheim, Norway

Tian Li
tian.li@risefr.no
RISE Fire Research, Trondheim, Norway

Pantelis Velanas
pvelanas@accelligence.tech
ACCELIGENCE LTD, Nicosia, Cyprus.

Katerina Margariti
katerina.margariti@accelligence.tech
ACCELIGENCE LTD, Nicosia, Cyprus.

John Soltatos
john.soldatos@gmail.com
Innov-Acts Ltd, Nicosia, Cyprus

Carmine Pascale
carmine.pascale@stress-scarl.it
STRESS S.c.ar.l., Napoli, Italy.

ZKI Fire Monitoring Service of DLR

ABSTRACT

The "ZKI Fire Monitoring System" of the German Aerospace Center (DLR) is making new research methods in the field of satellite-based wildfire detection easily accessible to academics and practitioners alike. The system provides up-to-date wildfire information for Europe and other pre-selected areas. Aiming at monitoring burnt areas and their spatial development in near-real time, a fully automated approach was developed (Nolde et al., 2020, Nolde et al., 2021) and implemented in a thematic processing chain. The "ZKI Fire Monitoring System" is operated by the Center for Satellite Based Crisis Information (ZKI) of DLR. To achieve several updates per day, mid-resolution satellite data are used. The daily research products show the perimeters of burnt areas, processed minutes after satellite imagery becomes available. Detected areas are iteratively and automatically refined over a period of 10 days as new satellite data becomes available. This enables continuous improvement of the derived fire perimeters by minimizing the influence of disturbing factors such as cloud cover, smoke plumes and shadows. Beside the burnt area perimeter, other attributes such as the fire severity, the area affected, as well as the exact detection times are included in the data. It is therefore possible to track the evolution of each burnt area in space and time.

The aim of the "ZKI Fire Monitoring System" is to provide burnt area analysis in near-real time, complementing existing services in Europe. The comparison and combination of different approaches, methods and sensors is vital for new scientific achievements and progress in this topic.

In addition to the processing component a public web application was developed (https://services.zki.dlr.de/fire/) to ease fast and interactive analysis and exploration of the generated research data products. In a first step the burnt area data is published as standardized OGC web service by the Earth Observation Center (EOC) Geoservice (Dengler et al., 2013; https://geoservice.dlr.de/) and afterwards integrated in a dedicated frontend user interface. There the detected wildfires can be shown on a dynamic web map, built with DLR's Environmental and Crisis Information Systems (UKIS) software components for geoscientific data visualization (Muehlbauer, n.d.; https://github.com/dlr-eoc/ukis-frontend-libraries). The web services themselves can be integrated in other geographical information systems as well. This makes it especially easy for researchers and practitioners to access the data and use it for a more detailed situational awareness overview in their own information systems.

Dengler, K., Heinen, T., Huber, A., Molch, K., & Mikusch, E. (2013). The EOC Geoservice: Standardized Access to Earth Observation Data Sets and Value Added Products. *Proceedings of PV, Frascati, Italy*.

Muehlbauer, M. (n.d.). UKIS – Environmental and Crisis Information Systems. German Aerospace Center, Earth Observation Center. Retrieved February 14, 2024, from https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-5413/10560_read-21914

Nolde, M., Plank, S., Riedlinger, T. (2020). An Adaptive and Extensible System for Satellite-Based, Large Scale Burnt Area Monitoring in Near-Real Time. *Remote Sensing*, 12, 2162. https://doi.org/10.3390/rs12132162

Nolde, M., Mueller, N., Strunz, G., Riedlinger, T. (2021). Assessment of Wildfire Activity Development Trends for Eastern Australia Using Multi-Sensor Earth Observation Data. *Remote Sensing*, 13, 4975.

https://doi.org/10.3390/rs13244975

PRESENTER/DEMONSTRATOR

Lucas Angermann*
lucas.angermann@dlr.de
DLR e.V.

Dr. Michael Nolde
michael.nolde@dlr.de
DLR e.V.

Martin Mühlbauer
martin.muehlbauer@dlr.de
DLR e.V.

Mathias Böck
mathias.boeck@dlr.de
DLR e.V.

Torsten Heinen
torsten.heinen@dlr.de
DLR e.V.

Dr. Torsten Riedlinger
torsten.riedlinger@dlr.de
DLR e.V.

Factors for Volunteer User Satisfaction and System Success in Warning Apps

ABSTRACT

The success of information systems in human computer interaction builds on the user's satisfaction with the technology or application. Warning and emergency response systems provide alerts and instruction to specific groups of individuals. As an emergency response system, it is not designed to improve productivity or quality of life in regular use cases, but it applies to exceptional cases – emergencies – only. In our work, we investigate this application area using a warning app as part of a fire detection system in which volunteer helpers are automatically alarmed based on the measurements of a sensor network. The primary goal of this socio-technical system is to support the protection of historical buildings, whereas individual benefits are secondary and realized through the collaboration in the protection of valuable assets. Consequently, user satisfaction must be viewed form a different perspective: While classical user satisfaction in information system success, technology acceptance or human computer interaction builds on the ease of use and benefits created by using a technology or information system, satisfaction in emergency response must consider how users perceive their safety and trust in the system and if the system's purpose itself corresponds to the user's aim: the protection of the building and assets inside.

The presented study is part of the BRAWA[1] project. We present findings from a study comprising three parts, a screening study on a variety of items and psychological variables, a main study based on the BRAWA app including an in-app questionnaire and a post questionnaire on experiences and ratings with 79 participants. Each participant received three alerts during a period of two weeks, informing the recipient about a potential fire in the building and the task to investigate the area. One set of participants additionally received reminders between the second and third alert on the purpose and functionality of the app. Another set received additional graphical aids illustrating the probability of the potential fire. A third set received more detailed information on potential causes for the fire based on measurements from the sensor network. Three preregistered hypotheses[2] investigated the relationship of reminders, graphical aids, and the amount of information on the perception and reaction of participants to the alerts.

In our findings we show that the Protective Action Decision Model (PADM) is a fitting model to explain user satisfaction and system's success in the context of emergency warning systems. We draw conclusions on the hypotheses and formulate suggestions for future app designs in this domain, e.g. that graphical aids increase the trustworthiness of information. Our findings will be submitted to the International Journal of Human-Computer Interaction. We want to present them as poster on the ISCRAM 2024 conference to engage with practitioners and experts in crisis management and information system design and discuss further insights into system developments.

[1] 'BRAWA: Kulturgut bewahren durch Helfermotivation und geringe Brandwahrscheinlichkeiten ', [BRAWA: Preserving Cultural Goods Through Helper Motivation and Early Fire Detection], https://www.brawa.ovgu.de/en/, This work was supported by the German Federal Ministry of Education and Research [grant number 13N15416].

[2] Preregistration available at https://aspredicted.org/blind.php?x=8MK_MBH

PRESENTER/DEMONSTRATOR

Mira von der Linde, M.Sc.*
mira.vonderlinde@uni-muenster.de
University of Münster, Department of Psychology ORCID
http://orcid.org/0000-0002-2384-9846

Dr. Michael Middelhoff
michael.middelhoff@wi.uni-muenster.de
University of Münster, Department of Information Systems ORCID
https://orcid.org/0000-0002-4167-9027

Destina Karadaş, B.Sc.
dkaradas@uni-muenster.de
University of Münster, Department of Psychology

Prof. Dr. Meinald T. Thielsch
thielsch@uni-muenster.de
University of Münster, Department of Psychology ORCID
http://orcid.org/0000-0001-8493-9071

Seismic Variables: Impact of Distance and Azimuth on Landslide Forecasting

ABSTRACT

In recent years, machine learning (ML) methods have gained popularity for predicting landslides following earthquakes, replacing traditional statistical or numerical simulation approaches. Most ML studies in this field relied on the Peak Ground Acceleration (PGA) to predict the landslides. Yet, the accuracy of PGA-based models can be compromised due to the limited availability of near-source observations and spatial information regarding seismic faults. This limitation can result in significant discrepancies between the accumulated PGA and the actual ground shaking, which cannot be overlooked. This study introduces a novel approach that utilizes ML models based on two seismic features: the distance and azimuth to the epicenter. The ML model using the distance and azimuth to the epicenter performs better than the model using PGA in terms of accuracy as well as actual prediction results. The findings indicate that these models exhibit superior performance in predicting the spatial distribution of landslides resulting from seismic events, compared to models relying solely on PGA. Furthermore, considering their computational efficiency and cost-effectiveness, the ML models utilizing distance and azimuth to the epicenter as inputs could be used as an alternative to PGA-based models for evaluating the impact of co-seismic landslides. Particularly, these models prove advantageous in near real-time scenarios and settings that require high spatial resolution.

PRESENTER/DEMONSTRATOR

*Zhenguo Zhang **
zhangzg@sustech.edu.cn
Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen, China

Yang Shi (to be at the conference first time)
Juanito_yangshi@163.com
Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen, China

Changhu Xue
chh_xue@163.com
State key laboratory of space-earth integrated information technology, Beijing 100095, China

Yu Feng
yu.feng@ngi.no
Department of offshore energy, Norwegian Geotechnical Institute, Sandakerveien 140, 0484 Oslo, Norway

Prototyping data-driven disaster mapping for multiple roles

ABSTRACT

Working on analog situation and crisis maps for major incidents is still common practice in Germany. Analog situation maps are updated manually. This is time-consuming and prone to errors. The time delay between receiving information and issuing commands is estimated to be over 90 minutes in some cases (cf. Lamers 2016[1]). It is not possible to work synchronously on a common situation picture regardless of location. The fact that more and more data is available also means that we find ourselves in a kind of unstructured jungle of information. This flood of data can only be quickly perceived and prioritized by one person to a limited extent.

Although there are attempts to convert these situation and crisis maps into a digital form, this usually only results in a simple 1:1 transfer of the analog map without exploiting the opportunities and possibilities of digital systems. The result is a digital system that enables location-independent and cross-device work but offers little added value in its handling compared to the analog form, although the available data (e.g. sensors, cameras, weather data, traffic data) is constantly increasing and can make a positive contribution to the interpretation of the situation map using artificial intelligence. Supporting applications with artificial intelligence, on the other hand, can quickly evaluate this data and alert the user to specific information in the event of anomalies.

In the SPELL research project, funded by the German Federal Ministry for economic affairs and climate action, the German Red Cross as a disaster response organization has the task of mediating between practice and development to communicate practical requirements and technological innovation opportunities between all parties involved.

In order to communicate these needs and innovation opportunities in a way that is understandable to all those involved, we at the German Red Cross are developing a visual design prototype of a practical and user-friendly digital situation map. Our guiding principle: as much recognition value as necessary, as much innovation as possible. The focus here is on the role-based networking of all management levels and emergency responders (e.g. political decision-makers, command staff, control center, ...) and the integration of AI services via the SPELL platform.

The demonstrator is based on familiar structures and symbols known from the situation map display. This recognition value is essential to ensure intuitive operation of the system by those involved, without the need for extensive training, and thus to achieve acceptance in practical applications. The demonstrator shows various applications of artificial intelligence developed as part of the SPELL

project and available to the operational and command staff.

The demonstrator allows access to the situation map by various levels of command. The views and applications adapt to the individual needs and access rights of the users.

PRESENTER/DEMONSTRATOR

Simon Franke*
s.franke@itc.drk.de
German Red Cross Rhineland Palatinate e. V. Information and Technology Center

Marcel Bender
m.bender@itc.drk.de
German Red Cross Rhineland Palatinate e. V. Information and Technology Center

[1]*Lamers, Christoph (2016): Stabsarbeit im Bevölkerungsschutz. Historie, Analyse und Vorschläge zur Optimierung. Edewecht: S+K Verlagsgesellschaft Stumpf + Kossendey mbH*

Alexander Rudolph
a.rudolph@itc.drk.de
German Red Cross Rhineland Palatinate e. V. Information and Technology Center

Luisa Schmidt
l.schmidt@itc.drk.de
German Red Cross Rhineland Palatinate e. V. Information and Technology Center