A Multi-Dimensional Analysis of User Classification, Sentiment, and Network Influence in Tornado-Focused Social Media Communication

Samiha Karim Subah
School of Civil Engineering & Environmental Science, University of Oklahoma
Email. Samiha.karim.subah-1@ou.edu

Arif Mohaimin Sadri
School of Civil Engineering & Environmental Science, University of Oklahoma
Email. sadri@ou.edu

ABSTRACT

During major disaster events, social media serves as a critical platform for user interactions, real-time updates and resource coordination. However, the rapid spread of disaster risk information complicates crisis communication. This study systematically categorized social media users based on their engagement behaviors. The study classifies users into Bot and Non-Bot (automated and human-operated accounts, respectively), with Non-Bot users further subcategorized as Individuals, Public and Private Agencies, Media and others. These categories were analyzed across sentiment, discussion topics, temporal activity patterns, and network interactions. Using advanced natural language processing and network science methods, this study examined geotagged tornado-related posts on X from Oklahoma between 2020 and 2022 and compared trends across years. Results show that social media communication is predominantly citizen-driven, with individuals occupying the most central network positions, media acting as secondary hubs, and bots exerting limited influence. Sentiment patterns further reveal differentiated stakeholder roles. These findings offer actionable insights for emergency managers to improve communication accuracy and coordination.

Keywords

Social media, risk communication, disasters, social networks.

INTRODUCTION AND MOTIVATION

Computing for disasters has the transformative potential to significantly strengthen preparedness and resilience by systematically collecting, integrating, and converting diverse data streams into secure, actionable intelligence for emergency officials. By enabling faster decision-making and more coordinated response efforts, it can help save lives, improve outcomes for affected populations, accelerate economic recovery, and even stimulate new technology-driven job sectors. Realizing this potential, however, demands interdisciplinary collaboration and sustained efforts to address challenges related to scope, scale, complexity, and uncertainty inherent in disaster environments (Consortium, 2012; Pu & Kitsuregawa, 2013). Social media platforms such as X, Facebook, and Instagram serve as important sources of real-time information during natural disasters and emergencies. Social media enables access to on-the-ground information, facilitates information sharing and crowdsourcing, and captures diverse responses from different user groups during disasters. Beyond traditional sources such as television and newspapers, social media enables the rapid retrieval, production, and dissemination of real-time information during emergencies, positioning it as a critical tool for disaster preparedness, warning, response, and recovery.

Crisis informatics has emerged as an interdisciplinary field that examines human behavior and response during disasters through the lens of social media data and increasingly pervasive information and communication technologies. The self-generating and rapidly shareable nature of social media substantially accelerates information production and diffusion, positioning these platforms as critical infrastructures for emergency communication by providing real-time insights into public sentiment, attention dynamics, and collective behavior. Example of such studies include the shooting in Virginia shooting, California wildfires (Hughes et al. 2008), earthquakes in China (Qu et al., 2011; Qu et al., 2009), Red River floods and Oklahoma grassfires (Vieweg et al., 2010) and Oklahoma tornado (Ukkusuri et al., 2014) among others. During Hurricane Sandy, social media served as a vital channel for information dissemination, particularly for residents in New York and New Jersey who faced power outage and limited access to traditional media such as radio and television. Through smartphones and online platforms, affected populations were able to receive timely updates and share situational information, underscoring the growing importance of digital communication networks during large-scale disasters.(Kaufman et al., 2012). Individuals were more likely to evacuate if they relied on social media for weather-related information during Sandy (Sadri et al., 2016).

X enables users to share short messages of up to 280 characters and follow other accounts, thereby forming large-scale interconnected structures that function simultaneously as social networks and information dissemination networks (Kryvasheyeu et al., 2016; Myers et al., 2014). From an emergency research perspective, many researchers used X to study the service characteristics (Guy et al., 2010; Li & Rao, 2010), retweeting activity (Kogan et al., 2015; Starbird & Palen, 2010), situational awareness (Power et al., 2014; Vieweg et al., 2010), online communication of emergency responders (Hughes et al., 2014; St Denis et al., 2014), text classification and event detection (Caragea et al., 2011; Earle et al., 2012; Imran et al., 2013; Kumar et al., 2014; Sakaki et al., 2010), devise sensor techniques for early awareness (Kryvasheyeu et al., 2015), quantifying human mobility (Wang & Taylor, 2014, 2015), and disaster relief efforts (Gao et al., 2011). Unlike survey-based datasets with limited geographical coverage and capacity to adequately capture risk communication, concerns (e.g., storm surge, wind speed) and needs (e.g., availability of gas, shelter locations); social media datasets are enriched with user activity information and geo-locations.

Crises inherently create a pressing need for authentic information, which often results in an influx of data from various sources (Qadir et al., 2016). As conditions continue to evolve, individuals and communities increasingly rely on social media platforms to access timely information and interpret shifting narrative contexts.(Momin et al., 2025).Thus social media analytics are increasingly employed to infer collective emotions and concerns, thereby supporting agencies' capacity to anticipate compliance challenges and tailor warning messages and protective guidance to evolving public responses (Ilyas & Sharifi, 2025). However, this growing reliance on social media data is accompanied by recognition that communication effectiveness cannot be evaluated solely in terms of message reach or engagement metrics. Instead, "soft" determinants particularly trust in institutions, source credibility, and perceived transparency consistently mediate whether risk messages are interpreted as legitimate and acted upon (Bonfanti et al., 2023). Complementing this, a recent scoping review of disaster risk communication determinants identifies trust-building practices, message clarity, and organizational credibility, alongside technological and institutional capacities, as core drivers of effective risk communication outcomes.(Mohsenzadeh et al., 2025). Therefore, to better understand the dynamics of information dissemination, it is essential to examine how users communicate over time, their behavioral patterns, and their individual roles in the propagation of information.

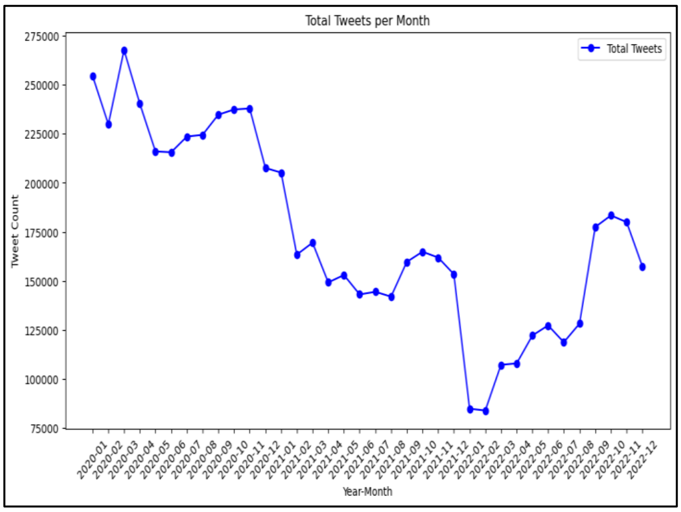
As such, in recent years, social media has become an indispensable tool during crises emergencies, enabling real-time information sharing, situational awareness, and public engagement (Sadri et al., 2017; Sadri et al., 2018). Social platforms such as X (formerly known as Twitter), Facebook, Reddit, Instagram serve not only as channels for official updates as well as spaces where individuals, organizations, and communities coordinate resources, express sentiments, and share firsthand experiences. X was selected due to its open-access API, real-time information flow, and transparent interaction structures (e.g., mentions, retweets), which enable large-scale network analysis; in contrast, Facebook and Instagram have more restrictive data access and limited visibility of interaction networks. Previous studies have examined the role of social media during tornado events, highlighting its importance for real-time situational awareness and information exchange among affected communities (Vieweg et al., 2010; Ukkusuri et al., 2014). These studies show that individuals actively share on-the-ground observations, warnings, and damage reports during tornado outbreaks, contributing to rapid information diffusion. However, existing research on tornado-related communication remains relatively limited compared to other hazards, particularly in terms of understanding stakeholder roles and network influence dynamics. However, the unregulated and rapid spread of information on social media poses significant challenges, such as the dissemination of misinformation, which can confuse the public, undermine trust, and obstruct timely crisis response efforts. Addressing these challenges requires a deeper understanding of the diverse actors involved in social media platforms. While previous research has explored content trends and platform dynamics, limited attention has been given to the systematic categorization of users and their behaviors over time(Beskow & Carley, 2019; Newell et al., 2021) . The goal of this study is to fill this gap in the literature and systematically examine stakeholder roles, communication patterns, and influence dynamics in tornado-related social media discourse in Oklahoma in order to improve understanding of crisis communication processes and support more effective disaster response strategies.

The specific research questions of this study are listed below:

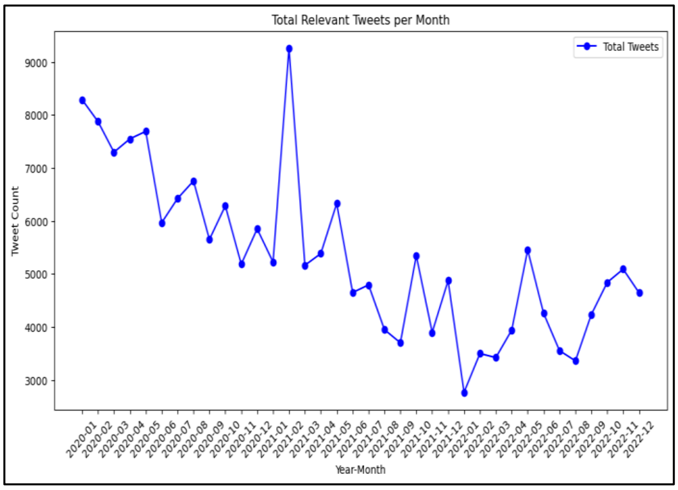
a) How do different stakeholder groups (individuals, agencies, media, and bots) differ in their communication behavior during tornado events?
b) What are the dominant sentiment patterns and discussion topics across stakeholder categories?
c) How is structural influence distributed within the tornado-related social media network, and which groups occupy central positions?
d) How consistent are stakeholder roles, influence patterns, and sentiment dynamics across multiple years?
e) What implications do these communication and influence patterns have for improving disaster response coordination and crisis communication strategies?

DATA COLLECTION AND DESCRIPTION

Social media platforms provide unique features to release data using an Application Programming Interface (API) (Sayce, 2019; X API, 2021). Using social media, users can post their concerns, needs, or reactions, and each data point can be stored as a tuple collected with the following information: user (name, unique identifier, location), text (description, time, unique identifier, latitude, longitude), count of followers, friends, favorites and statuses, user mentions, and others. This study adopts a multi-stage methodology to analyze social media user behavior during tornado disasters. The process consists of six interrelated steps: data collection and preprocessing, user classification, topic modeling, sentiment analyses, temporal analyses, and network analyses. A total of 6.2 million tweets with geo-locations and timestamps were collected from 2020 to 2022 using disaster-related keywords, Non-English, irrelevant (tweets were identified as those lacking substantive disaster-related content, including off-topic discussions, metaphorical uses of hazard-related terms), and duplicate tweets were filtered out. Preprocessing included text cleaning, tokenization, and basic metadata extraction (e.g., timestamps, user location, and engagement metrics). The tweets are analyzed by calculating various metrics like total tweets per month, total unique users, total unique users tweeting each month and total mentioned users per month and visualized them. A list of disaster-related keywords is defined to filter tweets containing these words in the text column. This filtering allows for a more focused analysis of tweets relevant to weather events and disasters. Once the relevant tweets are filtered, the metrics like total tweets per month, total unique users tweeting each month and total mentioned users per month are observed, visualized and compared to original tweet data. Data preprocessing steps also included the removal of non-English tweets, duplicates, spam, and bot-generated content where possible. Retweets and quote tweets were retained to analyze interaction dynamics.



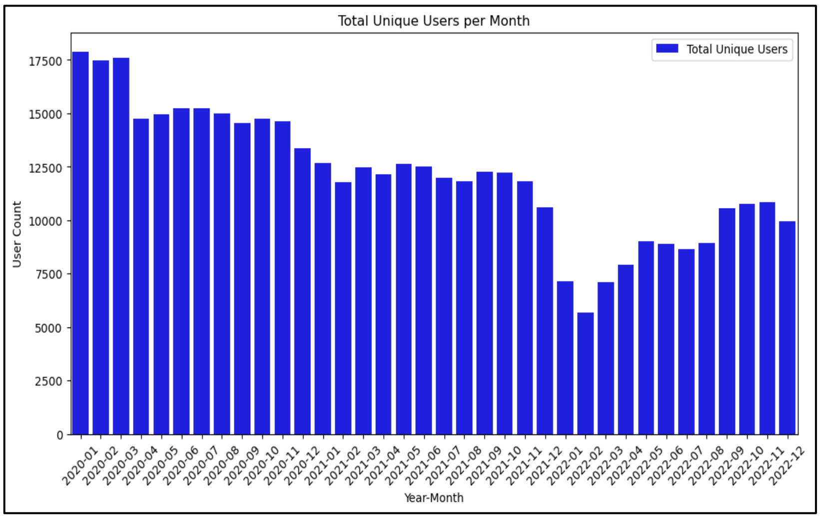
(a)



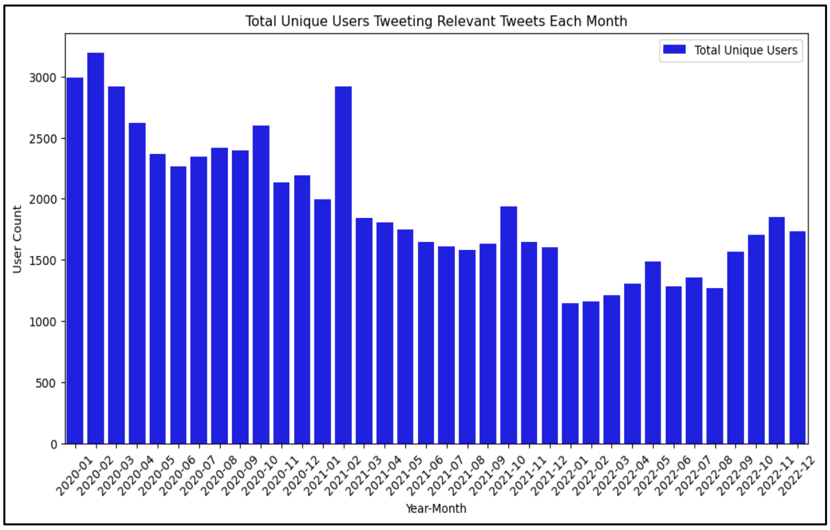
(b)

Figure 1. Number of total (a) and relevant (b) tweets per month (2000-2022)

Figure *1*shows that total monthly tweet volume was high in early 2020, declined steadily through 2021, dropped sharply around early 2022, and then partially rebounded in late 2022 before fluctuating downward again. Figure 1(b) shows a similar but more variable pattern for relevant disaster related tweets, with an overall downward trend from 2020 to 2022, punctuated by occasional spikes and a modest recovery toward the end of the period. The high tweet volumes in early 2020 likely reflect intense public attention during major severe weather outbreaks in Oklahoma, while the overall decline through 2021–2022 may indicate disaster fatigue, overlapping crises (e.g., winter storms, flooding, and broader national emergencies), and shifting public focus, followed by renewed spikes corresponding to particularly impactful tornado or storm events that temporarily reactivated public concern and social media engagement.



(a)



(b)

Figure 2. Number of total (a) and relevant (b) X users per month (2000-2022)

Figure *2* shows that the total number of unique users discussing overall social media topics in Oklahoma was highest in early 2020, steadily declined through 2021, dropped sharply around early 2022, and then gradually recovered. These trends indicate reduced but stabilizing public participation over time. Figure 2(b) mirrors this pattern for users posting disaster related topics, with an overall downward trend and brief spikes likely tied to major storm events, followed by a modest rebound as significant disasters renewed focused engagement.

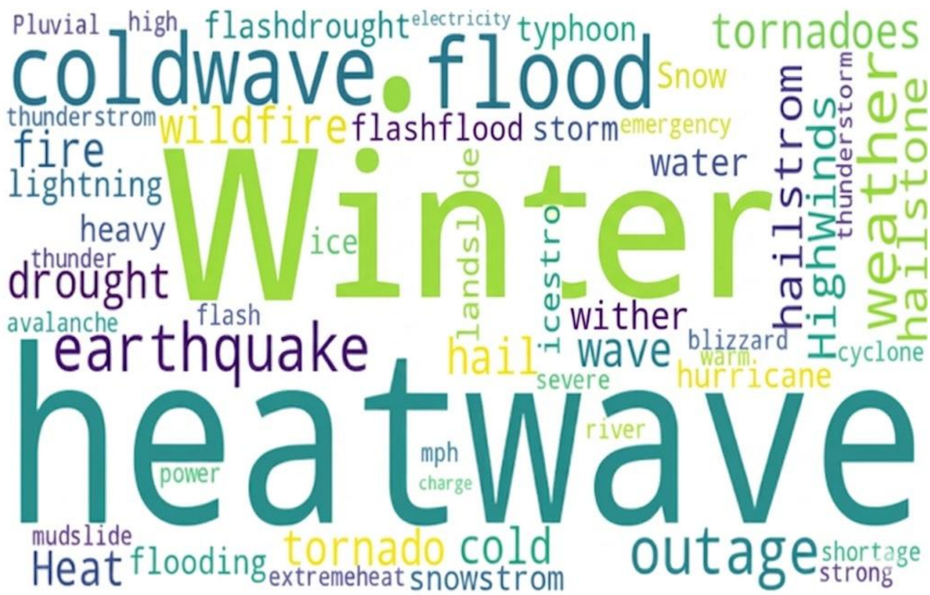


Figure 3. Keyword Frequency of natural hazard discussions in social media for Oklahoma (2000-2022)

The word cloud shown in **Figure** *3*is dominated by winter, heat wave, cold wave, and flood, indicating that temperature extremes and flooding are the most prominent hazards in the dataset. Other frequently appearing terms such as tornadoes, wildfire, hailstorms, earthquakes, and hurricanes indicate a broad range of disaster types being discussed. Overall, the visualization highlights strong attention to extreme weather events in Oklahoma, particularly seasonal temperature extremes and water-related hazards.

METHODOLOGIES

User Classification

To identify bot accounts, we employed a supervised machine learning approach using behavioral and profile-based features, including retweet frequency, tweeting intervals, followee–follower ratio, mention patterns, and profile metadata. These features capture common bot-like behaviors such as high posting frequency, automated interaction patterns, and abnormal network characteristics. Multiple classifiers (e.g., Random Forest, SVM, and Gradient Boosting) were evaluated using cross-validation to ensure robust classification performance.(Badr et al., 2024; Friedman, 2000). Despite applying this classification framework, bots were found to have limited influence in tornado-related communication. This contrasts with prior studies in other disaster and crisis contexts, where bots have played a more prominent role in amplifying misinformation and shaping public discourse. The relatively low bot influence observed in this study suggests that tornado-related communication in Oklahoma is driven more by organic, citizen-led interactions rather than automated amplification. Further, the data is then split into **features** (e.g., retweet count, follower–following ratio, average tweeting interval, mention count, and profile description length) and **labels** (bot = 1, non-bot = 0), where the labels represent the ground-truth classification of each user. Once the data is prepared, it is ready to be used for training the machine learning model, which learns to distinguish between bots and non-bots based on the features provided. Named Entity Recognition was applied to the user profile name, keyword filtering using the user profile description, detection of missing values (NaN) in profile fields, and the length of the user profile description (measured as the number of characters in the bio). Missing values (NaN, Not a Number) in profile fields were also identified and handled during preprocessing. Once the individuals are identified by above steps and separated, the agencies are further classified into nonprofit, public, media and private organizations using relevant keywords filtering. Each type of user is properly labelled in the dataset under user classification (Hagen et al., 2020).

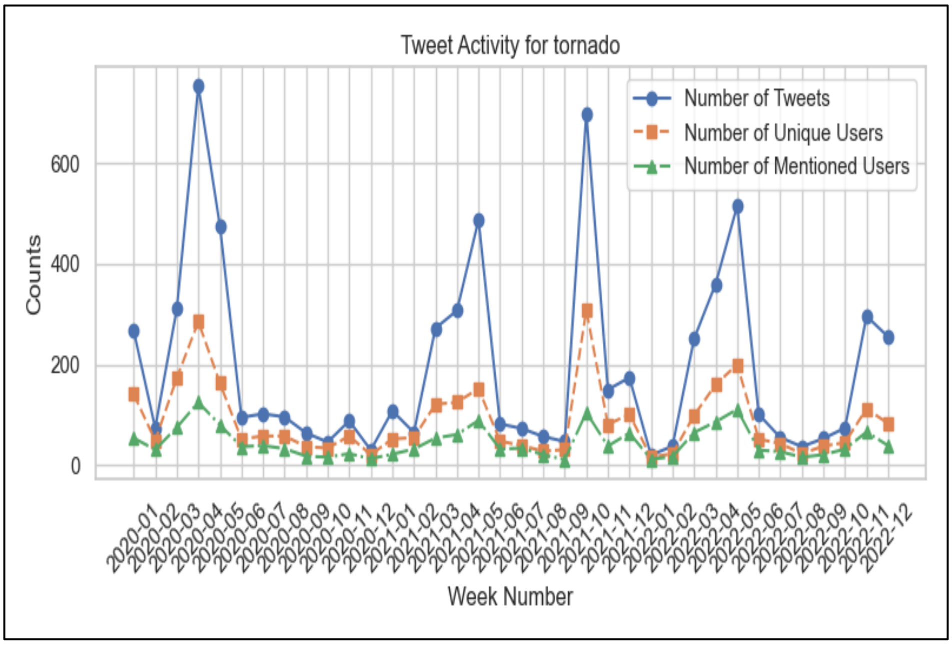
Disaster Type Classification and Topic Modeling

The data were further prepared for training a Bidirectional Encoder Representations from Transformers (BERT) model, a pre-trained deep learning language model designed to capture contextual relationships in text. The fine-tuned BERT model was used for natural hazard classification, categorizing tweets into ten distinct disaster-related hazard classes. The initial dataset contains varying numbers of samples for each disaster type, making it imbalanced. To address this issue, the target sample size for each hazard category was determined based on the total number of available instances. For rare hazard types (e.g., wildfire), all available samples were retained to preserve limited information, while for more frequent hazard categories, proportional sampling was applied to obtain representative subsets. This sampling strategy was adopted to mitigate class imbalance, reduce model bias toward majority classes, and improve classification performance across underrepresented disaster types(Alam et al., 2021). For other disaster types, proportional sampling is applied to achieve a representative sample size. This step ensures that each disaster type contributes appropriately to the model training process, preventing the model from being biased towards the more frequent disaster categories. After sampling, the dataset was shuffled to randomize the order of the samples (Mao et al., 2024). This helps reduce the chances of bias in model training and ensures a more generalized learning process. To train the BERT model for disaster classification, a pre-trained BERT model was considered and a classification layer tailored to disaster categories was added. The tokenized and padded input data is fed into the model, and training is performed using cross-entropy loss and the AdamW. optimizer(Ludwig et al., 2021). The model is fine-tuned over several epochs using weighted loss techniques to address any remaining data imbalance. After training, the model is evaluated on a test set using metrics like accuracy, precision, recall, and F1-score, with a confusion matrix providing further insights into possible misclassifications. Topic modeling was performed on tornado-related tweets, focusing on the tornado hazard category(Zhao et al., 2011) . Latent Dirichlet Allocation (LDA) is used to identify key topics discussed during the tornado disasters(Wang & Ma, 2024). The coherence score is calculated for different numbers of topics to find the optimal model(Wu et al., 2020). After determining the optimal number of topics, the LDA model is trained, and topics are labeled based on their content (e.g., "Personal Experiences," "Severe Weather Conditions", "Recover Efforts" among others).

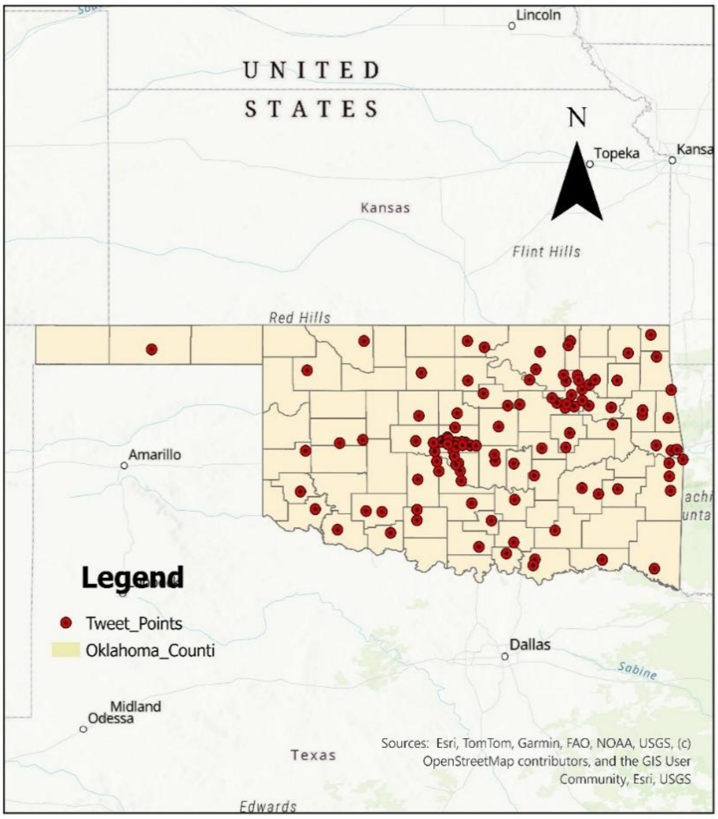
Network Analyses

We analyzed the dynamics of communication within user-mention networks formed around tornado-related discussions. A user-mention network was constructed by mapping interactions between users who mention one another in tweets, where each node represents a user and each directed edge represents a mention, indicating the flow of information from one user to another. This network structure was used to examine how information flows and how different user types engage during disaster events(Al Momin et al., 2025). In-degree centrality was used to quantify the level of attention received by each user, while PageRank centrality was employed to capture influence based on both the number and importance of incoming connections(Freeman, 2004; Page et al., 1999)

RESULTS



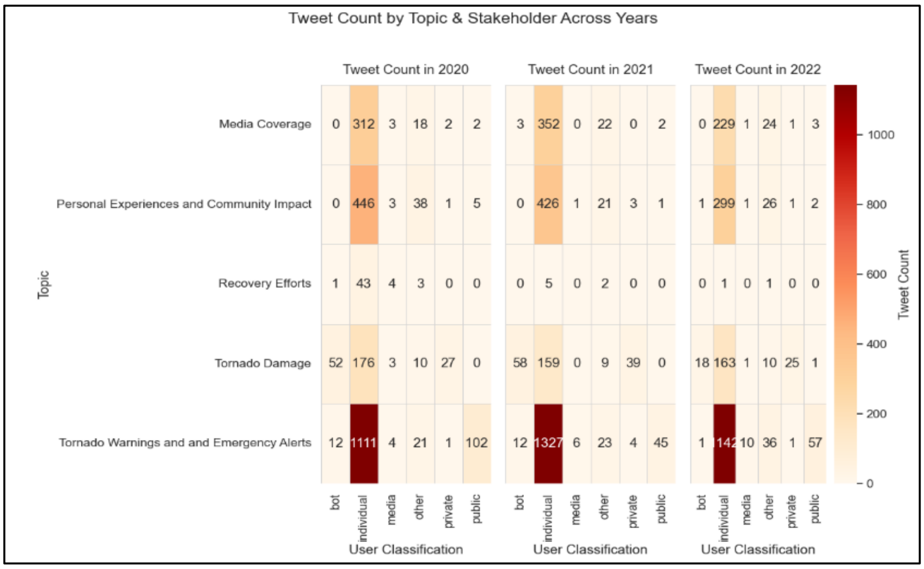
(a)



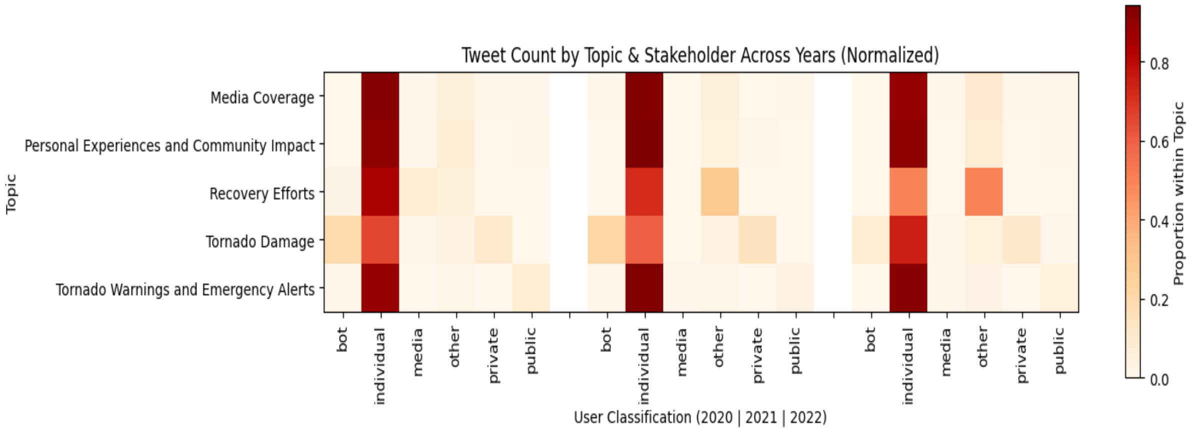
(b)

Figure 4. (a) Seasonal patterns in tornado-focused X Activity in Oklahoma (2000-2022). (b) Spatial patterns of tornado related tweets in Oklahoma (2000-2022)

Figure ***4*** shows clear seasonal spikes in tornado-related social media activity on X that align with Oklahoma's primary tornado season in spring (typically April–June) and a secondary peak in late fall. During these periods, the number of tweets, unique users, and mentioned users all rise sharply, reflecting heightened public attention, information sharing, and engagement with meteorologists and emergency agencies during active severe weather outbreaks. In contrast, activity drops substantially during off-season months, consistent with Oklahoma's lower tornado risk outside peak periods.



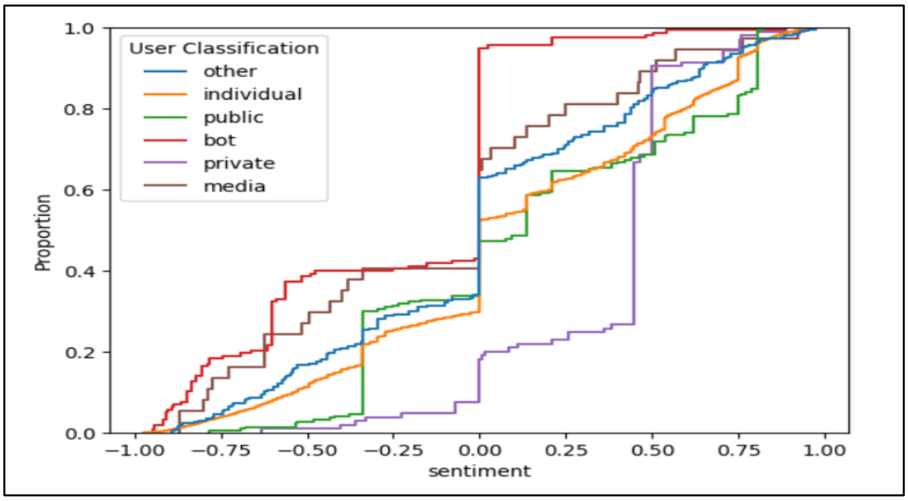
(a)



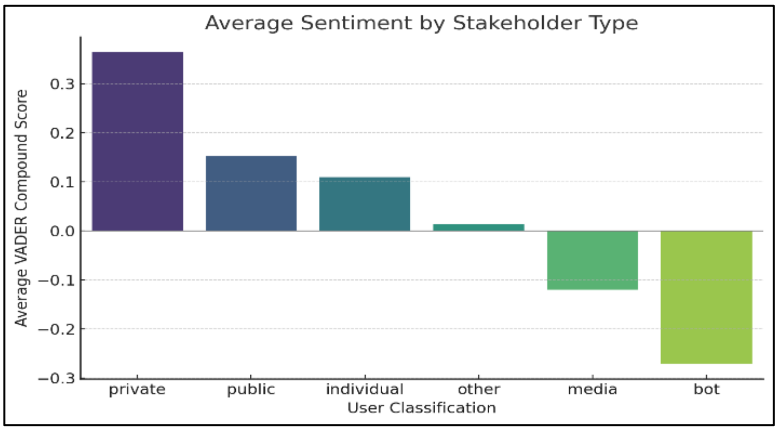
(b)

Figure 5. Distribution of tornado focused tweet topics across stakeholder groups: (a) direct tweet counts; (b) normalized distribution

Heatmaps in ***Figure 5*** present the direct count and normalized distribution of tornado focused tweet topics across stakeholder groups (bots, individuals, media, public, private, and others) from 2020 to 2022, illustrating the proportion of content each group contributes within each thematic category. Across all years, individual users consistently dominate discourse related to media coverage, personal experiences, tornado warnings, and damage, suggesting that disaster communication on X is primarily citizen driven. Media accounts show comparatively concentrated contributions in damage- and recovery-related topics, while bots maintain a relatively limited and stable presence in covering tornado damage, indicating that automated amplification plays a secondary role relative to organic public engagement. It was also observed that users classified into public, private and other categories had very minimal contribution related to tornado related crisis communication. Overall, the figure highlights persistent stakeholder role differentiation in tornado-related communication over time.



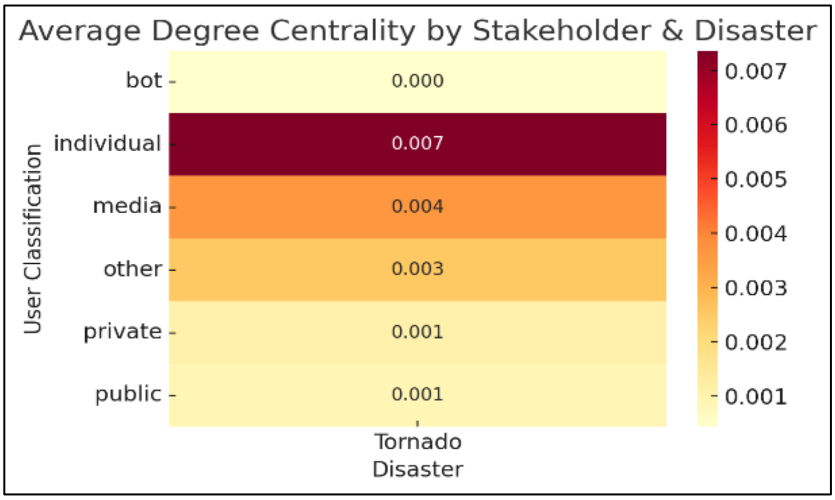
(a)



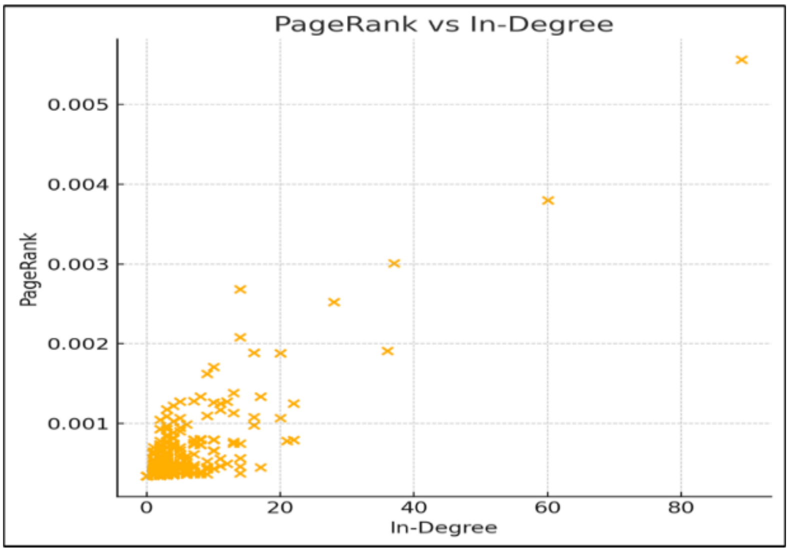
(b)

Figure 6. Sentiment analysis of tornado-related social media posts by user classification: (a) Cumulative sentiment distribution by user type during tornado discussions and (b) Average VADER compound sentiment scores across stakeholder types

The cumulative distribution plot in **Figure *6***(a) shows that individual users as well as private and public organizational accounts tend to exhibit more positively skewed sentiment distributions, with a greater proportion of tweets falling in neutral-to-positive ranges, whereas traditional media and bot accounts are more concentrated toward neutral or negative sentiment. The bar chart in Figure 6(b) of average VADER compound sentiment reinforces this interpretation: private and public users display the highest average positive sentiment scores, individuals remain moderately positive, media accounts trend slightly negative, and bots exhibit the most negative average sentiment.



(a)



(b)

Figure 7. Network centrality analysis of users during tornado events. (a) Heatmap of average degree centrality by stakeholder type, showing which user groups are more connected. (b) Scatter plot of PageRank versus in-degree for individual users

The network analyses conducted in this study also collectively characterize the structural influence and information flow within the Oklahoma tornado–related risk communication in social media. For example, the heatmap of average degree centrality in Figure 7(a) indicates that individual users occupy the most central positions in the communication network, followed by media accounts, while bots, private, and public institutional accounts exhibit comparatively low centrality. This suggests that tornado-related discussions are primarily structured around highly connected individuals, likely local residents, weather enthusiasts, and community voices, rather than being dominated by institutional or automated actors. Media accounts maintain moderate centrality, reflecting their continued relevance as information hubs, but they do not appear to monopolize network connectivity. The PageRank versus in-degree scatter plot in Figure 7(b) further supports this interpretation by demonstrating a positive relationship between in-degree (number of incoming connections) and PageRank (network influence), indicating that accounts receiving more mentions or retweets tend to accumulate greater structural authority.

DISCUSSIONS

The seasonal patterns indicate that tornado-related social media activity closely aligns with hazard occurrence in Oklahoma. Peaks during spring and late fall reflect increased public awareness and engagement, highlighting social media's role in real-time information exchange. Higher activity during these periods suggests intensified communication among individuals, meteorologists, and emergency agencies, while lower activity in off-season months reflects reduced public attention. Overall, these findings emphasize the event-driven nature of crisis communication on social media.

The results indicate that tornado-related communication on X is predominantly citizen-driven, with individuals leading discussions across most topics. Media accounts play a more focused role in covering damage and recovery, while bots have limited influence, suggesting that automated amplification is secondary to organic engagement. The minimal contribution from public and private entities highlights a potential gap in institutional participation. Overall, these patterns reflect clear and consistent differentiation in stakeholder roles over time.

In the context of Oklahoma's recurrent tornado seasons, this divergence in **Figure *6*** likely reflects functional role differentiation: citizens using social media for community support and recovery narratives, media emphasizing impact and hazard severity, and bots amplifying alert- or damage-focused content. Overall, the findings suggest that tornado-related social media risk communication in Oklahoma is not uniformly negative but instead characterized by emotionally layered communication shaped by stakeholder function.

However, the distribution is highly skewed, with a small number of nodes exhibiting disproportionately high in-degree and PageRank values consistent with a hub-like structure typical of crisis communication networks. In the context of Oklahoma tornado events, this pattern suggests that while many users participate, influence is concentrated among a limited set of highly visible actors, likely key individuals and media sources who serve as focal points for information dissemination during severe weather.

The network analysis highlights that tornado-related communication on social media is largely driven by individuals, who occupy the most central and influential positions in the network. This suggests that information flow is primarily shaped by highly connected users, such as local residents and community voices, rather than institutional or automated actors. Media accounts retain an important but secondary role as information hubs, while bots and institutional accounts exhibit relatively limited influence. The positive relationship between in-degree and PageRank further indicates that users who receive more mentions or retweets gain greater visibility and authority within the network. The application of network analysis metrics such as degree centrality and PageRank provides practical insights into identifying influential users and understanding information flow during tornado events. High-centrality users can be leveraged by emergency agencies to disseminate critical information more effectively, while network structures can help identify key communication hubs and potential gaps in information reach. Additionally, examining network patterns can support the design of targeted communication strategies to improve coordination and enhance the overall effectiveness of disaster risk communication.

CONCLUSIONS AND FUTURE SCOPE

This work-in-progress study finds that Oklahoma tornado–related social media risk communication is highly seasonal and event-driven, with sharp spikes in activity during peak tornado periods and declining engagement over time punctuated by disaster-specific surges. Communication is predominantly citizen-driven, as individual users occupy the most central and influential network positions, while media accounts serve important, but secondary hub roles and bots contribute comparatively limited structural influence. Sentiment patterns further reveal differentiated stakeholder functions: individuals as well as public and private agencies express relatively more positive expressions, whereas media and automated accounts express negative connotations. Future research should extend this framework to multiple disaster types (e.g., floods, winter storms, wildfires) to compare cross-hazard communication dynamics, incorporate spatial proximity to disaster locations to assess how geographic exposure shapes engagement and sentiment, and examine influence mechanisms more deeply including information cascade dynamics, influence pathways, and temporal shifts in network centrality among others to better understand how information authority and public response evolve during compound and concurrent hazard events. Future research should also incorporate more granular user and topic classification frameworks to generate high-resolution insights into stakeholder roles, narrative evolution, and micro-level communication dynamics across disaster phases. Emergency management agencies should engage with highly connected individual users, as they play a key role in information dissemination during tornado events. Partnering with media organizations can further enhance the spread of accurate and timely information. The findings also suggest prioritizing human-centered communication strategies, such as real-time updates and community engagement, to build trust and improve response effectiveness. Additionally, the limited presence of public and private institutions highlights the need for more proactive and consistent institutional communication on social media during disasters.

ACKNOWLEDGMENTS

The research presented in this paper was supported by the National Science Foundation under Grant No. CAREER-2339100. Any opinions, findings, conclusions, and recommendations expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Al Momin, K., Sadri, A. M., Olofsson, K., Muraleetharan, K., & Gladwin, H. (2025). Information switching patterns of risk communication in social media during disasters. *IEEE Transactions on Big Data*.

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. https://doi.org/10.1609/icwsm.v15i1.18115

Badr, H., Awahida, Z., Essgaer, M., Agaal, A., & Ahessin, A. (2024). *Named Entity Recognition for Identifying Entities Related to Illegal Migration in Libya: An Analysis of Twitter Textual Data*. https://doi.org/10.1109/MI-STA61267.2024.10599731

Beskow, D., & Carley, K. (2019). *Agent Based Simulation of Bot Disinformation Maneuvers in Twitter*. https://doi.org/10.1109/WSC40007.2019.9004942

Bonfanti, R. C., Oberti, B., Ravazzoli, E., Rinaldi, A., Ruggieri, S., & Schimmenti, A. (2023). The Role of Trust in Disaster Risk Reduction: A Critical Review. *International Journal of Environmental Research and Public Health*, *21*(1), 29. https://doi.org/10.3390/ijerph21010029

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324

Caragea, C., McNeese, N., Jaiswal, A., Traylor, G., Kim, H.-W., Mitra, P., Wu, D., Tapia, A. H., Giles, L., & Jansen, B. J. (2011). Classifying text messages for the haiti earthquake. Proceedings of the 8th international conference on information systems for crisis response and management (ISCRAM2011),

Consortium, C. C. (2012). Computing for Disasters: A Report from the Community Workshop. In.

Earle, P. S., Bowden, D. C., & Guy, M. (2012). Twitter earthquake detection: earthquake monitoring in a social world. *Annals of Geophysics*, *54*(6).

Freeman, L. (2004). The Development of Social Network Analysis.

Friedman, J. (2000). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, *29*. https://doi.org/10.1214/aos/1013203451

Gao, H., Barbier, G., Goolsby, R., & Zeng, D. (2011). *Harnessing the crowdsourcing power of social media for disaster relief.*

Guy, M., Earle, P., Ostrum, C., Gruchalla, K., & Horvath, S. (2010). Integration and dissemination of citizen reported and seismically derived earthquake information via social network technologies. International Symposium on Intelligent Data Analysis,

Hagen, L., Neely, S., Scharf, R., & Keller, T. E. (2020). Social Media Use for Crisis and Emergency Risk Communications During the Zika Health Crisis. *Digital Government: Research and Practice*, *1*(2), 1–21. https://doi.org/10.1145/3372021

Hughes, A. L., St Denis, L. A., Palen, L., & Anderson, K. M. (2014). Online public communications by police & fire services during the 2012 Hurricane Sandy. Proceedings of the 32nd annual ACM conference on Human factors in computing systems,

Ilyas, B., & Sharifi, A. (2025). A systematic review of social media-based sentiment analysis in disaster risk management. *International Journal of Disaster Risk Reduction*, *123*, 105487. https://doi.org/10.1016/j.ijdrr.2025.105487

Imran, M., Elbassuoni, S. M., Castillo, C., Diaz, F., & Meier, P. (2013). Extracting information nuggets from disaster-related messages in social media. *Proc. of ISCRAM, Baden-Baden, Germany*.

Kaufman, S., Qing, C., Levenson, N., & Hanson, M. (2012). Transportation during and after Hurricane Sandy. 1–34.

Kogan, M., Palen, L., & Anderson, K. M. (2015). Think local, retweet global: Retweeting by the geographically-vulnerable during Hurricane Sandy. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Kryvasheyeu, Y., Chen, H., Moro, E., Van Hentenryck, P., & Cebrian, M. (2015). Performance of social network sensors during Hurricane Sandy. *PLoS one*, *10*(2), e0117288.

Kryvasheyeu, Y., Chen, H., Obradovich, N., Moro, E., Van Hentenryck, P., Fowler, J., & Cebrian, M. (2016). Rapid assessment of disaster damage using social media activity. *Science advances*, *2*(3), e1500779.

Kumar, S., Hu, X., & Liu, H. (2014). A behavior analytics approach to identifying tweets from crisis regions. Proceedings of the 25th ACM conference on Hypertext and social media,

Li, J., & Rao, H. R. (2010). Twitter as a rapid response news service: An exploration in the context of the 2008 China earthquake. *The Electronic Journal of Information Systems in Developing Countries*, *42*.

Ludwig, S., Mayer, C., Hansen, C., Eilers, K., & Brandt, S. (2021). Automated Essay Scoring Using Transformer Models. *Psych*, *3*(4), 897–915. https://doi.org/10.3390/psych3040056

Mao, Y., Yan, X., Hu, Z., Zhang, X., Ye, Y., & Yu, H. (2024). Memory-aware continual learning with multi-modal social media streams for unsupervised disaster classification. *Advanced Engineering Informatics*, *62*, 102654. https://doi.org/10.1016/j.aei.2024.102654

Mohsenzadeh, Y., Farrokhi, M., Saatchi, M., Pisheh, E. G., & Khankeh, H. (2025). A scoping review of factors affecting disaster risk communication, 2000-2023. *Journal of Education and Health Promotion*, *14*(1). https://doi.org/10.4103/jehp.jehp_1099_24

Momin, K. A., Sadri, A. M., Olofsson, K., Muraleetharan, K. K., & Gladwin, H. (2025). Information Switching Patterns of Risk Communication in Social Media During Disasters. *IEEE Transactions on Big Data*, *11*(4), 1733–1744. https://doi.org/10.1109/tbdata.2024.3524828

Myers, S. A., Sharma, A., Gupta, P., & Lin, J. (2014). Information network or social network?: the structure of the twitter follow graph. Proceedings of the 23rd International Conference on World Wide Web,

Newell, E., Jurgens, D., Saleem, H., Vala, H., Sassine, J., Armstrong, C., & Ruths, D. (2021). User Migration in Online Social Networks: A Case Study on Reddit During a Period of Community Unrest. *Proceedings of the International AAAI Conference on Web and Social Media*, *10*, 279–288. https://doi.org/10.1609/icwsm.v10i1.14750

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank Citation Ranking : Bringing Order to the Web. The Web Conference,

Power, R., Robinson, B., Colton, J., & Cameron, M. (2014). Emergency situation awareness: Twitter case studies. International Conference on Information Systems for Crisis Response and Management in Mediterranean Countries,

Pu, C., & Kitsuregawa, M. (2013). Big data and disaster management: a report from the jst/nsf joint workshop. *Georgia Institute of Technology, CERCS*.

Qadir, J., Ali, A., Ur Rasool, R., Zwitter, A., Sathiaseelan, A., & Crowcroft, J. (2016). Crisis analytics: big data-driven crisis response. *Journal of International Humanitarian Action*, *1*(1). https://doi.org/10.1186/s41018-016-0013-9

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,

Qu, Y., Wu, P. F., & Wang, X. (2009). Online community response to major disaster: A study of Tianya forum in the 2008 Sichuan earthquake. System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on,

Sadri, A. M., Hasan, S., Ukkusuri, S. V., & Cebrian, M. (2017). Crisis Communication Patterns in Social Media during Hurricane Sandy. *Transportation Research Record*, 0361198118773896.

Sadri, A. M., Hasan, S., Ukkusuri, S. V., & Lopez, J. E. S. (2018). Analysis of social interaction network properties and growth on Twitter. *Social Network Analysis and Mining*, *8*(1), 56.

Sadri, A. M., Ukkusuri, S. V., & Gladwin, H. (2016). The Role of Social Networks and Information Sources on Hurriacne Evacuation Decision Making. *Natural Hazards Review (accepted)*.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. Proceedings of the 19th international conference on World wide web,

Sayce, D. (2019). *The Number of tweets per day in 2019*. Retrieved 01/16/2020 from https://www.dsayce.com/social-media/tweets-day/

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. 11th International ISCRAM Conference,

Starbird, K., & Palen, L. (2010). *Pass it on?: Retweeting in mass emergency*. International Community on Information Systems for Crisis Response and Management.

Ukkusuri, S., Zhan, X., Sadri, A. M., & Ye, Q. (2014). Use of social media data to explore crisis informatics: Study of 2013 Oklahoma tornado. *Transportation Research Record: Journal of the Transportation Research Board*(2459), 110–118.

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,

Wang, Q., & Ma, B. (2024). Enhancing BERTopic with Pre-Clustered Knowledge: Reducing Feature Sparsity in Short Text Topic Modeling. *Journal of Data Analysis and Information Processing*, *12*(04), 597–611. https://doi.org/10.4236/jdaip.2024.124032

Wang, Q., & Taylor, J. E. (2014). Quantifying human mobility perturbation and resilience in Hurricane Sandy. *PLoS one*, *9*(11), e112608.

Wang, Q., & Taylor, J. E. (2015). Resilience of human mobility under the influence of typhoons. *Procedia Engineering*, *118*, 942–949.

Wu, J., Son, G., & Wang, S. (2020). A Competency Mining Method Based on Latent Dirichlet Allocation (LDA) Model. *Journal of Physics: Conference Series*, *1682*, 012059. https://doi.org/10.1088/1742-6596/1682/1/012059

X API. (2021). *Real-time access to the global conversation, right at your fingertips.* Retrieved 02/06/2024 from https://developer.twitter.com/en/products/twitter-api

Zhao, W., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011). *Comparing Twitter and Traditional Media Using Topic Models* (Vol. 6611/2011). https://doi.org/10.1007/978-3-642-20161-5_34