

# Supporting the management of spontaneous volunteers through AI-generated task descriptions

**Michael Klafft**

Jade University of Applied Sciences and Fraunhofer FOKUS

[Michael.Klafft@jade-hs.de](mailto:Michael.Klafft@jade-hs.de)

## ABSTRACT

Since about 2015, smartphones apps have been used in some areas of Germany to coordinate first responders in case of emergencies, and crises. Initially, the focus of such applications were medical emergencies and sudden cardiac arrests in particular. Later, use cases for crowdtasking apps have been broadened in order to address all types of disaster scenarios, and tasks. Additionally, disaster and emergency managers use such tools nowadays to coordinate untrained, spontaneous volunteers without specific qualifications in times of crises. While those volunteers can be of valuable support for some disaster management tasks, such as establishing emergency shelters, emergency management professionals have sometimes perceived the management of spontaneous helpers as burdensome in stressful situations. Furthermore, in times of crises, some emergency managers also felt overwhelmed by the situation and would have needed more guidance. Artificial intelligence could provide tools and features to support emergency managers in the process of managing spontaneous volunteers. This paper presents some initial results from a desktop exercise where generative AI was used to support the creation of task descriptions for spontaneous volunteers, and discusses future uses of AI with regard to volunteer management in disaster situations.

## Keywords

Spontaneous volunteers, crowdtasking system, artificial intelligence, crisis communication, desktop exercise.

## INTRODUCTION

Bidirectional crisis communication between citizens and authorities in times of crises has been a challenge: This is because a limited number of emergency management professionals have to communicate with a large number of citizens, making it difficult to consider the individual perspective outside of high-priority emergency calls. This paper discusses actual uses of artificial intelligence in the overall communication process between emergency managers and citizens. Thereafter, an addition to a crowdtasking tool for spontaneous volunteers is presented which uses generative artificial intelligence to assist emergency managers in the formulation of detailed task descriptions. Finally, it is described how the tool was tested in a desktop exercise, and strengths and weaknesses of the results are discussed.

## PREVIOUS RESEARCH

Artificial intelligence is one approach to improve bidirectional communication between emergency managers and citizens in times of crises. Previous works have looked into how artificial intelligence can be used to extract situational information from texts and images posted by citizens on social media, with a particular focus on microblogs (for an overview, see Ogie et al., 2018). Use cases were the identification of tweet locations (Kiatpanont et al., 2016), the identification of affected or vulnerable areas (Panday & Natarajan, 2016), and thematic classifications of postings (e.g., if they contain information on injured or missing persons, infrastructural damage, or emergency shelters, Anbalagan & Valliyammai, 2016). Another area where artificial intelligence has been applied in crisis communication is sentiment analysis, in order to identify situations of panic, sorrows and their prevalence within a population, or the degree of polarization of the public (Beigi et al., 2016). Furthermore, AI-based sentiment analysis has been used to determine the impact of alert messages on a population (Brynielsson et al., 2014). All these use cases have in common that they only indirectly support the communication between authorities and citizens, or facilitate the flow of information between the latter. However, some applications have been developed that directly support communication with the citizens. One such example are chatbots, which

were provided by some authorities during the COVID-19 pandemic. These were used to distribute health-related information, to assess a person's individual infection risk, to combat the distribution of fake news, or for monitoring of symptoms of those people who had been infected (Almalki & Azeez, 2020). Sautter et al. (2021) proposed the use of artificial intelligence tools as translation assistants in cases where a person calls an emergency operations center in a foreign language, and first implementations of such systems already exist (Liedtke, 2024). In this paper, the author provides an example of using generative AI in emergency messaging – here in particular task descriptions and calls for action issued to volunteers in times of emergencies and crises. Previous research in this field was conducted by Miller (2024) who analyzed the usefulness of generative AI in emergency messaging in general. Key findings were that generative AI can increase efficiency and improve message clarity, but that there are concerns among emergency managers about accuracy, context, and cultural sensitivity. However, emergency managers tend to evaluate emergency messaging systems using generative AI much more positively after actually using them (Miller, 2024). Although Miller's works show that some research on the use of generative AI in emergency messaging is already ongoing, such technologies have so far not yet been used in practice to support the direct communication between disaster managers and spontaneous volunteers in times of crisis.

Crowdtasking of volunteers in times of crisis, has, however, long been supported by information technology in general. Earlier approaches focused on web-based tools like Ushahidi which had already been used in 2010 for crowdsourced crisis mapping after an earthquake in Haiti (Norheim-Hagtun & Meier, 2010). More recently different types of crowdtasking apps have been developed to support the assignment of tasks to volunteers in times of emergencies or crises (for an overview, see e.g., Fuchs-Kittowski, 2021). Examples from Germany include "Mobile Retter" (Stroop et al., 2015) or "Katretter". The latter system is currently operational in Berlin, and has, for example, been used over 16000 times between October 2020 and 2022 to activate first responders in case of out-of-hospital cardiac arrests. While generally open for all types of volunteers, the overwhelming majority (98%) of activated volunteers for this specific use case did have some medical knowledge or expertise (Pommerenke et al., 2023). However, non-professional volunteers have also been activated for more general tasks requiring less previous knowledge such as constructing emergency shelters when needed.

The proposed combination of a generative AI tool with a crowdtasking tool to create task descriptions for (mostly unskilled) volunteers has the potential to support emergency communicators in times of high workloads and stress levels. AI generated texts provide additional guidance to the operators. In contrast to predefined messages used in emergency communication today, generative AI is more flexible (since the AI component can also adapt to unforeseen situations). AI generated descriptions also allow to create detailed instructions and advice in a matter of seconds, thus potentially saving time in critical situations.

## SET-UP FOR THE DESKTOP EXERCISE

The prototype of the crowdtasking exercise used a clone of the Katretter system (Katretter, 2025) as the background infrastructure for activating volunteers. However, the following changes were applied to the system: a dedicated front-end was used to facilitate displaying and modifying the AI-generated task descriptions. These task descriptions were generated by the ChatGPT 3.5 generative AI tool which was embedded via its API. As soon as the operator had created the overall task title, such as "constructing emergency shelters", three requests were sent to the generative AI tool. These requests were as follows<sup>1</sup>:

- Create a description of the assignment for a deployment with the following title: `${this.taskTitle}`. The description should come from the point of view of an emergency operations center and be addressed to volunteers. The description should roughly explain what the assignment is about, so that volunteers know what to expect. The description should not be longer than 5 sentences. The individual sentences should not be numbered.
- Formulate up to five different instructions or tasks for a deployment with the following title: `${this.taskTitle}`. The instructions should come from the point of view of an emergency operations center and be directed at volunteers. The instructions should describe exactly what volunteers have to do when they arrive at the scene. The individual instructions or tasks should be numbered. Only the key points are needed, without introduction or farewell. The individual instructions or tasks should not have a heading or a hyphen.
- Formulate up to three different pieces of advice for an operation with the following title: `${this.taskTitle}`. The advice should come from the point of view of an emergency operations center and be addressed to volunteers. The pieces of advice should give volunteers additional information about what they need at the disaster scene. For example, what equipment they should bring. Only the key points

---

<sup>1</sup> Please note that requests were made in German. What is presented here, however, is the English translation of these requests.

are needed, without an introduction or farewell. The individual instructions should not have a heading or a hyphen.

Thus, the generative AI returned three types of information: a detailed and extended task description, up to five instructions, and up to three different additional advices. Please note that “\${this.taskTitle}” is the placeholder for the overall task title entered by the operator.

The desktop exercise then focused on a relevant scenario where spontaneous volunteers can provide added value for emergency management. This scenario was identified in discussions with emergency managers from the city of Cottbus in Eastern Germany. The scenario focused on establishing emergency shelters, which has repeatedly been a pressing necessity in recent years mostly due to arrivals of a large numbers of refugees (e.g., from Syria or Ukraine). However, establishing emergency shelters also is a relevant issue in all disaster scenarios where people have to be evacuated over an extended period of time. The overall tasks used in the desktop exercise were “setting up tents and beds”, then a more specific instruction “setting up camp beds”, and then a more “social” instruction, specifically “looking after children at an emergency shelter”. Each query was submitted to the AI component several times, in order to check the consistency and reliability of the answers provided. Furthermore, it was tested if the responses reached a number of test users who simulated professional volunteers in the exercise.

## EVALUATION SCHEME FOR AI-GENERATED TASK DESCRIPTIONS

The evaluation scheme for the generated task-description is based upon previous findings from the literature. As Miller (2024) states, the following criteria are particularly relevant for AI generated crisis messaging: message clarity, context, accuracy, and cultural sensitivity. Additional criteria were derived from information technology acceptance modeling (TAM), and are related to the key determinants of acceptance: usefulness and ease of use (Davis, 1985). AI generated task descriptions are useful if they are – in addition to the criteria mentioned above – complete, consistent, easy to understand, and provide an adequate level of detail (i.e., enough information for recipients to understand what they need to do, and to assess whether they are able to fulfil the requested task). Task descriptions are also easy to use if no or only very minor corrections are needed. Thus, the complete set of evaluation criteria is as follows:

- Accuracy
- Consistency
- Message clarity
- Context
- Cultural sensitivity
- Completeness
- Understandability
- Adequate level of detail
- Degree of necessary corrections

The detailed assessment of the generated messages using these criteria was conducted by the researcher himself ex-post after the desktop exercise.

## RESULTS AND DISCUSSION OF THE DESKTOP EXERCISE

As can be expected from using generative AI, results varied between different test runs. In all cases, the overall task description and call for action were reasonable. In the instruction and advice sections, however, some sentences in the responses were not so suitable or even unsuitable and would have to be corrected or deleted by the operator. These issues will be discussed in more detail below:

In one instance, the AI provided answers in English, although the query was made in German – thus not meeting the understandability criterion. This weakness was later addressed by explicitly instructing the AI to provide responses only in German. The wording used by the generative AI also fluctuated between more formal language and informal language (here, the use of the German word “Du” which usually is only used among friends, young people, and colleagues and the more formal term “Sie”). The following examples (all translated from German) illustrate the findings above:

### Task 1: Setting up tents and beds

In this use case, one more concrete **task description** read as follows:

“Our assignment “constructing tents and beds” focuses on preparing accommodation for arriving people. You will work in teams to set up tents, assemble beds, and arrange other necessary equipment. The project requires both

physical labor and team spirit, which is why we are counting on your active support. We will provide you with meals and professional instruction so that you can fully concentrate on the task at hand. Together we will make an important contribution to supporting those in need on site”.

Interestingly, the generative AI turned to use motivational speak like “we will make an important contribution” despite not having been specifically instructed to do so.

The corresponding **set of instructions** was:

“Please report to the operations management at the central information point and register as a helper. Divide yourselves into groups of at least four people and choose a group leader to coordinate setting up the tent. Start unloading the tent materials from the transport vehicles and check that all the materials are present using the lists provided. Assemble the tents using the enclosed assembly instructions and make sure that all tents are set up securely and stably. Prepare the bedrooms with mattresses and bedding by cleaning the respective areas and setting up the beds according to the designated room arrangement.”

These instructions need to be discussed in more detail. It is highly sensible to register as volunteers / helpers before joining the disaster management effort, for example in order to assure that they are insured during the planned activities. This instruction therefore make sense. Whether there is an information point for helpers in any case is, however, not clear – here the operator may have to adapt the proposed wording, thus having a potential negative impact on the criterion “Degree of necessary corrections”. The instruction to assemble the tents using the enclosed instructions is also somewhat contradictory to the promise from the task description that helpers would receive “professional instructions” on site, thus the criterion of consistency was violated to some degree. The generative AI also focused on the first part of the task (constructing tents) and mentions the second part (setting up beds) only briefly. This is most likely due to the fact that the number of instruction sentences was limited to just five. Figure 1 shows how the proposed instructions were presented to the operator. As can be seen, the operator can easily delete sentences that are deemed as unsuitable by clicking on the red cross after each sentence. Furthermore, it is also possible to freely edit the proposed text before sending.

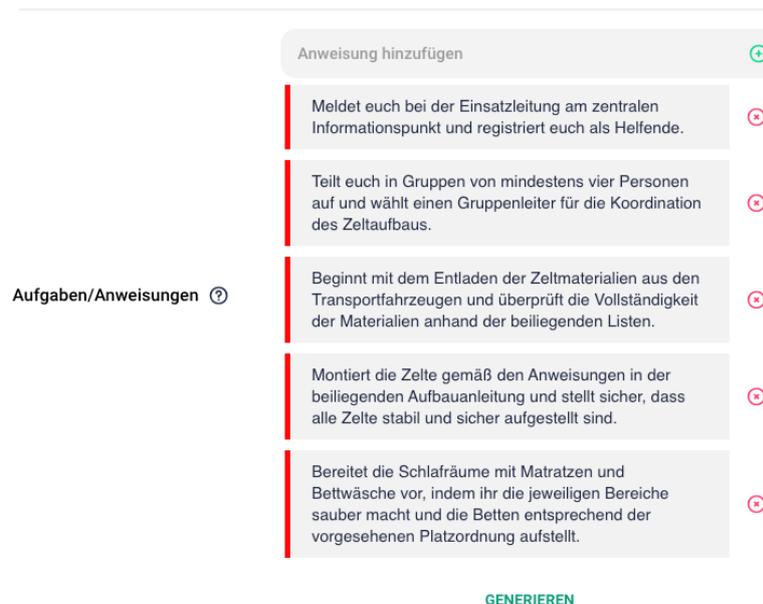


Fig. 1: Instructions for the task “constructing tents and beds” (in German) with the option to edit or delete

As can also be seen, the operator can restart the instruction generation if the results are not suitable at all (which during testing and trials would, for example, have been necessary in the case where instructions were returned in English, hoping that a second attempt delivers answers in German).

The **additional advice** included the following items:

“Please bring weatherproof clothing and sturdy shoes to keep you comfortable in all weather conditions. Please remember to bring gloves and possibly safety goggles for handling tools. Don't forget snacks and drinks to keep you going during the operation.”

While overall sensible, sturdy shoes are good but in fact not the best option. If available, dedicated safety footwear is even better and might even be more relevant from a work safety perspective in this case than safety goggles, so the instruction was to some degree incomplete. The wording is also unrealistically positive (“comfortable”), and

again the last sentence is in contradiction to the general task description which states that food will be provided. However, this could also be due to the fact that the three queries were submitted independently by the crowdtasking software and therefore the AI may not be aware that they were related to the same case.

### Task 2: Setting up camp beds

Since the first task was very broadly formulated, instructions for setting up beds were rather brief and vague. A more specific task title (setting up tents and beds) led to more specific instructions and advice, but also led to some hallucinating (thus potentially violating the aspect of correctness). Therefore, much more double-checking and editing on the operator's side would have been needed (having a negative impact on the degree of necessary corrections). The last four out of five instructions were very specific most likely related to one type of tent beds and may not be universally applicable. The **instructions** were as follows:

“When you arrive at the operation site, take one camp bed each. To set up the camp bed, place the side supports and the floor cover parallel to each other on the surface provided. Connect the side supports from top to bottom through the center struts. Insert the tent poles into the holes in the side supports and secure them with the top braces and clamps. Finally, you can place the floor covering on the side supports to complete the assembly of the camp bed”.

In addition to being very specific and thus most likely applicable to just one type of bed (if not completely fictitious), these task instructions are also much too detailed for a call to action (inadequate level of detail). The extended use of technical language might also frighten away spontaneous volunteers (lack of understandability). Likewise, the **additional advice** given was not too helpful:

“Bring an impact wrench and a set of bits for the mounting screws. Depending on the weather, it is recommended that you bring a blanket and sleeping bag. Pack a smart, portable screwdriver and sturdy replacement screws in case a screw goes missing.”

Here, it remains unclear what a “smart” screwdriver is (lack of understandability) and how it might be helpful (a normal portable screwdriver might just be fine). The order to bring “replacement screws” with you is also only moderately helpful (lack of context) without given more specific information on what types of screws might be needed. An impact wrench might also be too much for setting up camp beds unless these shall be firmly connected to the floor or walls (a possible violation of the criterion of correctness).

### Task 3: Looking after children at an emergency shelter

While the first two types of tasks were more manual / technical, the third task rather had a social focus and asked volunteers to help to look after children at an emergency shelter. The following example shows that the texts created by the generative AI for this specific case were by and large reasonable. The detailed **call for action** read as follows:

“As a volunteer, you will look after children and young people in an emergency shelter. Your role is to support the work of the educators and to provide psychosocial support to young people when needed. Together you create a safe environment for the children to develop. Help us to improve the quality of life for children in emergency shelters.”

Although only some volunteers may have the qualification to provide psychological support, a large number of volunteers should have the skills to provide social support during the crisis. So this call to action is by and large suitable. In addition to this call for action, the following set of **instructions** was provided:

“Go to the shelter and register with the people on site. Support the shelter staff by taking care of the children. Play with the children, tell stories and share time and attention with them. Engage with the children, help with homework or accompany them during their free time. Your commitment will help to ease the delicate situation for the children and reduce their psychological burden.” These instructions are reasonable. An adaptation may however be necessary if no schooling can be provided during longer stays at the emergency shelter (degree of necessary corrections / lack of context). Registration is particularly relevant in this case, not only for insurance purposes, but also to control access to the children and to give the managers on site the possibility to reassign or reject volunteers deemed to be not suitable for supporting childcare.

The **additional advice** in this case read as follow:

“Please bring sufficient personal protective equipment (gloves, face mask, etc.) and wear it during your assignment. Please also bring toys, children's books and comfortable clothing to support the children in the emergency shelter as much as possible. Be mindful and respectful of children so that they feel safe.”

As can be seen, advice was strongly influenced by the pandemic situations. If the emergency management operators deem personal protective equipment not necessary, or if such equipment is available on site, the

corresponding sentence can be easily deleted. Although possibly relevant in this case, the cultural context was not taken into account by the AI in this call for action. Additional advice could / should also include an appeal to volunteers to take cultural or religious peculiarities / requirements of the children into account, e.g., with regard to food ingredients or preparation.

## CONCLUSIONS AND OUTLOOK

This paper provided first insights into potential uses of generative AI in crowdtasking management. Initial results suggest that – while the generated task descriptions by and large meet the defined quality criteria, the criterion of cultural sensitivity was completely overlooked by the generative AI. Some degree of corrections of proposed messages were also needed due to occasional violations of the understandability, consistency, and the completeness criterion. In some instances, too much level of detail was provided, and the context was not always adequately taken into account. Therefore, some more prompt engineering is still needed so that instructions and advice provided by generative AI tools become even more suitable for the crowdtasking of volunteers. While some issues like receiving responses in an unsuitable language have already been corrected by adding more instructions, some other issues observed need more research. There are, for example, some indications that queries must be designed in a way that they are not too vague on the one hand, but not too specific on the other hand. Too vague descriptions increase the likelihood to receive answers that are of less practical value. Too specific descriptions on the other hand, increase the likelihood of hallucinating or receiving too specific or technical answers which cannot be applied in practice or are likely to deter volunteers from participating. In addition to prompt engineering, comparing and evaluating the performance of different generative AI tools and models with regard to the use case discussed are another field of future research. Furthermore, systematic acceptance studies with both volunteers and emergency managers are needed – for an iterative improvement of the existing tool, and to make sure that emergency managers will accept AI support in the use case presented after practical experience, in line with Miller’s (2024) broader positive assessment of AI acceptance in emergency messaging.

## ACKNOWLEDGEMENTS

The author thanks Dustin Reimann and David Holz for the technical implementation of the extension of the component to include AI-generated task descriptions into an existing crowdtasking tool. The research work was in part funded by the German Federal Ministry of Economic Affairs and Climate Action through the SPELL project grant (FKZ 01MK21005B)

## REFERENCES

- Almalki, M., & Azeez, F. (2020). Health chatbots for fighting COVID-19: a scoping review. *Acta Informatica Medica*, 28(4), 241.
- Anbalagan, B., & Valliyammai, C. (2016). ChennaiFloods: leveraging human and machine learning for crisis mapping during disasters using social media. In *2016 IEEE 23rd international conference on high performance computing workshops (HiPCW)* (pp. 50-59). IEEE.
- Beigi, G., Hu, X., Maciejewski, R., & Liu, H. (2016). An overview of sentiment analysis in social media and its applications in disaster relief. *Sentiment analysis and ontology engineering: An environment of computational intelligence*, 313-340.
- Brynielsson, J., Johansson, F., Jonsson, C., & Westling, A. (2014). Emotion classification of social media posts for estimating people’s reactions to communicated alert messages during crises. *Security Informatics*, 3, 1-11.
- Davis, F. D. (1985). A technology acceptance model for empirically testing new end-user information systems: Theory and results. Diss. Massachusetts Institute of Technology, 1985.
- Fuchs-Kittowski, F. (2021). Mobiles Crowdsourcing zur Einbindung freiwilliger Helfer. In Reuter, C. (Ed.): *Sicherheitskritische Mensch-Computer-Interaktion: Interaktive Technologien und Soziale Medien im Krisen- und Sicherheitsmanagement* (pp. 581-604). Wiesbaden: Springer Fachmedien.
- Katretter (2025): Katretter – jede Hilfe hilft! [www.katretter.de](http://www.katretter.de) (accessed February 20<sup>th</sup>, 2025)
- Kiatpanont, R., Tanlamai, U., & Chongstitvatana, P. (2016). Extraction of actionable information from crowdsourced disaster data. *Journal of emergency management*, 14(6), 377-390.
- Liedtke, S. (2024). Vfdb: Mehr Einsatz von KI in Leitstellen, URL: <https://www.feuerwehrmagazin.de/wissen/vfdb-mehr-einsatz-von-ki-in-leitstellen-126178> (accessed February 16<sup>th</sup>, 2025).

- Miller, S. (2024). Emerging Technology in Emergency Management: Using generative artificial intelligence (AI) to improve emergency messaging. URL: <https://doi.org/10.13140/RG.2.21739.25124>.
- Norheim-Hagtun, I., & Meier, P. (2010). Crowdsourcing for crisis mapping in Haiti. *Innovations: Technology/Governance/Globalization*, 5(4), 81-89.
- Ogie, R. I., Rho, J. C., & Clarke, R. J. (2018). Artificial intelligence in disaster risk communication: A systematic literature review. In *2018 5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)* (pp. 1-8). IEEE.
- Pandey, N., & Natarajan, S. (2016). How social media can contribute during disaster events? Case study of Chennai floods 2015. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1352-1356). IEEE.
- Pommerenke, C., Poloczek, S., Breuer, F., Wolff, J., & Dahmen, J. (2023). Automated and app-based activation of first responders for prehospital cardiac arrest: an analysis of 16.500 activations of the KATRETTTER system in Berlin. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, 31(1), 105.
- Sautter, J., Ziegler, D., Drews, P., Münzner, A., Kirchner, J. (2021): Neue Technologien und sicherheitskritische Mensch-Technik-Interaktion in der Rettungskette: Plädoyer für eine frühe Berücksichtigung von Innovationshemmnissen in der Entwicklung. Mensch und Computer 2021 – Workshopband, Beitrag 370.