Improving Social Media Geolocation for Disaster Response by Using Text From Images and ChatGPT

Hafiz Budi Firmansyah
University of Geneva
Switzerland
hafiz.firmansyah@unige.ch

Valerio Lorini
Joint Research Center European Commission
Ispra, Italy

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

Jose Luis Fernandez-Marquez
University of Geneva
Switzerland
joseluis.fernandez@unige.ch

ABSTRACT

Social media can serve as a valuable source of timely and valuable information about the impacts on people and infrastructure during a disaster. However, due to the lack of geographical data in most social media posts, this information is often underutilized by first responders. This paper proposes and analyses an approach that combines text from social media posts with textual information extracted from images to improve the geolocation of social media data during a given disaster. The implementation incorporates ChatGPT in location prediction. We use real-world dataset from Twitter that represent four different events, including floods and earthquake. The experimental results demonstrate that our proposal improves the location prediction's quantity and precision. We expect that our findings help policymakers consider the application of the proposed methodology in disaster response.

KEYWORDS

ChatGPT, Geolocation, Social Media, Disaster Response

ACM Reference Format:
Hafiz Budi Firmansyah, Valerio Lorini, Mehmet Oguz Mulayim, Jorge Gomes, and Jose Luis Fernandez-Marquez. 2024. Improving Social Media Geolocation for Disaster Response by Using Text From Images and ChatGPT. In *2024 The 11th Multidisciplinary International Social Networks Conference (MISNC 2024), August 21–23, 2024, Bali, Indonesia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3675669.3675696

MISNC 2024, August 21–23, 2024, Bali, Indonesia

ACM ISBN 979-8-4007-1755-0/24/08
https://doi.org/10.1145/3675669.3675696

1 INTRODUCTION

During the disasters, on online social media platforms, users in affected areas post updates on broken infrastructure, injured victims, damage severity, and real-time situation of disaster sites [2]. From a policymaker's perspective, this rapid information flow might lead to effective emergency measures [11]. Social media posts may contain geolocation information that could prove critical for disaster responders (i.e., fire department, police, or search and rescue team) to assess the damaged site and respond to the situation immediately. However, the fact that only about 3% of social media posts include organized geolocation data [3, 10] presents obstacles to leveraging social media content in disaster response efforts.

The possibility of enhancing disaster response through social media content has attracted researchers to integrate natural language processing (NLP) methods and gazetteers to retrieve geolocation information. Many contributions in geolocating social media content are centered on text analysis, such as examining the text associated with images to determine the images' capture locations. These methods have demonstrated their ability to accurately identify geolocation with a significant percentage of social media posts at both country and city levels. Furthermore, text-based geolocation techniques have been enhanced through the integration of crowdsourcing to refine the precision of location. Nevertheless, while the integration strategy demonstrates high precision, the manual process of geolocation can delay the creation of disaster maps. Consequently, there is a need to augment the quantity of images geolocated automatically to reduce the time required for generating an accurate disaster area map.

In a social media post, it is common for users to attach images. The text within these images, including street names, traffic signs, and business or landmark names, serves a vital part in crowdsourcing image geolocations. Yet, to the best of our knowledge, in geolocation retrieval, automatically extracting text from social media images and combining this information with the message text in posts has not been previously investigated.

The study comprehensively investigates the value of the text embedded in images attached to social media posts for enhancing information quality during a disaster.

The remainder of the paper is structured as follows: Section 2 discusses related work. Section 3 presents the dataset and proposed approach. Section 4 shows the experimental results using real-world social media data from four disaster events. Section 5 discusses the findings. Finally, Section 6 reports the conclusion and future works.

2 RELATED WORK

The research on advanced machine learning and social media data for disaster response has been trending over the last decade. Latest advancements in technology have significantly enhanced the analysis of both text and images from social media. A recent study on text focused on understanding rescue requests on Hurricane Harvey where most of the requests came from the affected community [19]. Yigitcanlar et al investigated the role of Twitter users in Australia during disasters [17]. Their research provides policymakers with an approach to examining the geographical distribution, frequency of occurrence, and impact of various disasters through the analysis of tweets. A disaster map was created using non-authoritative data from social media posts during the Venice floods [13]. The work managed to estimate the floods' extent immediately by integrating social media data and digital surface models.

Within disaster management, besides using text, images from social media are also increasingly being utilized for classifying information and predicting locations. Numerous studies have emphasized and utilized the significance of images in the context of crisis events, including [1, 4, 7, 9]. Nguyen et al leveraged social media images to perform damage assessment [14]. Their experimental result demonstrated that advanced machine learning models can classify damage severity during disasters. Firmansyah et al introduced a pipeline that uses text appearing in social media images to determine their locations [8]. The study demonstrated reasonably good predictions at the country level.

In recent years, researcher implements machine learning to help in determining location in the disaster [5]. The challenge to get precise location is also applicable to efforts aimed at geolocating social media posts by combining machine learning and crowdsourcing [15]. In another study, Wu et al conducted an analysis based on multiple data resources, combining social media data, economic losses, and geo-information [16]. They found that the majority of tweets contain information about the hurricane or its impact with negative sentiment. The study demonstrated that there was a significant correlation between Twitter activity, disaster-related topics, and impacted areas.

Our study adopts a different approach compared to prior studies. Instead of using text or images, we combined both of them by extracting text from images and incorporating the extracted text into social media posts. To determine the location, we incorporated Large Language Model (LLM) in our approach

3 PROPOSED APPROACH AND DATASET

In this section, we introduce our proposed approach, which aims to address the challenges of identifying location from social media content through an innovative process. Additionally, we describe the dataset employed to validate our approach, carefully selected to ensure a comprehensive evaluation.

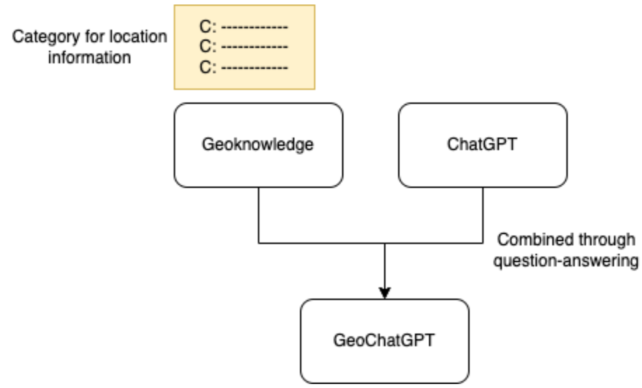

Figure 1: The fusion process of Geoknowledge and ChatGPT

3.1 Proposed approach

In this study, we investigate whether combining extracted text from an image with text improves location prediction using Large Language Models. To extract the text from the image automatically, we leveraged Object Character Recognition (OCR). Furthermore, we describe the details of the text extraction from social media images in the dataset subsection. After extracting the text from images using OCR, we concatenate the extracted text with the post text, separating both texts with a space. For predicting the location, we leverage the ChatGPT[1] model as LLM implementation. For generalization purposes, in this experiment, we used the freely available ChatGPT version 3.5. We also incorporate GeoChatGPT to obtain more precise location prediction [18]. GeoChatGPT fuses geo-knowledge of location descriptions and ChatGPT. The GeoChatGPT is trained through a question-and-answer mechanism (see Figure 1). Initially, geoknowledge provides different categories of the location followed by the location name. ChatGPT is trained by providing questions and answers. The training process uses a small number of data (less than 10 data). Finally, the fusion can predict precisely the location name and description from social media. Figure 2 illustrates the architecture of the proposed approach.

The concatenation of text from post with text from image aims to provide more information to ChatGPT. The input consists of two sets of text: one from a social media post and another extracted from an image, separated by a space.

ChatGPT is designed as a general-purpose chatbot that can answer relatively all questions from users. Since it is not specifically designed to predict a location, we leverage GeoChatGPT which is mentioned previously. The GeoChatGPT is trained through prompts by asking a question and providing its answer.

3.2 Dataset

To evaluate our proposed approach, we used Twitter data related to two disaster types: earthquakes and floods, utilizing both text and image modalities. This dataset was developed with the aid of the Social Media for Disaster Risk Management (SMDRM) Platform [12], as collected by the Copernicus Emergency Management Service [2].

[1] https://chat.openai.com/
[2] https://emergency.copernicus.eu/

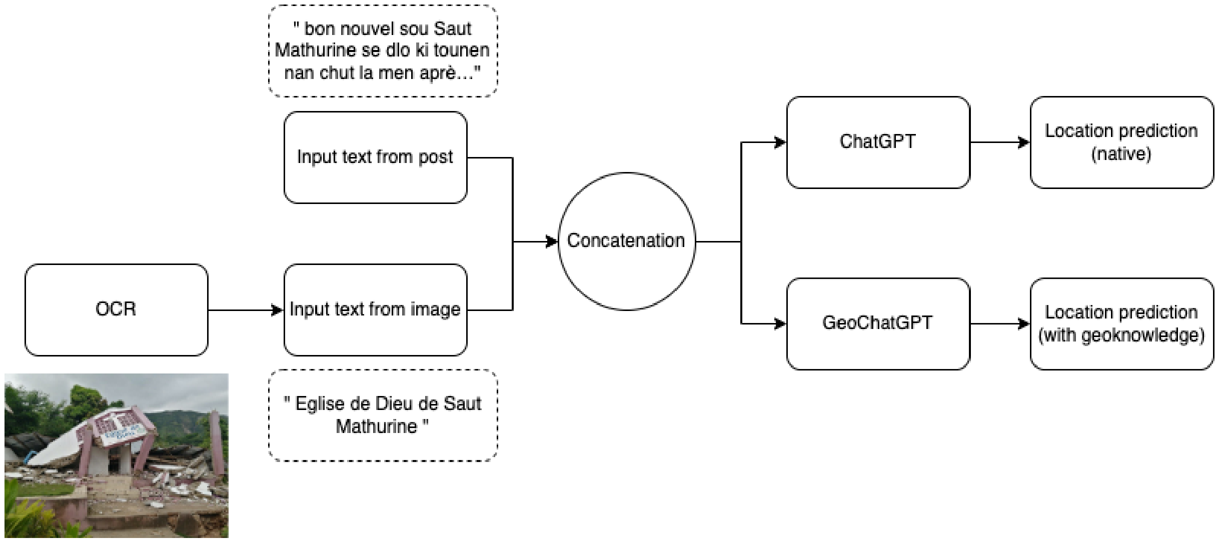

Figure 2: Architecture of proposed approach

It includes 5,430 text posts, out of which 4,714 were identified as unique. Additionally, the dataset encompasses 1,752 images across four disaster events: the Catania floods of 2021 (70 images), the Central European floods of 2021 (214 images), the Croatia earthquake of 2020 (549 images), and the Haiti earthquake of 2010 (961 images).

Out of the 1,752 images analyzed, 1,130 were found to contain text. To extract this text, we utilized the readily available technology provided by Amazon, namely, the Amazon Rekognition service [3]. Out of 1,130 images, Amazon Rekognition accurately identified text in 1,029 images. Subsequently, we refined this subset to 532 images by filtering out those with banners, news logos, image captions, and watermarks, focusing only on images containing genuine text.

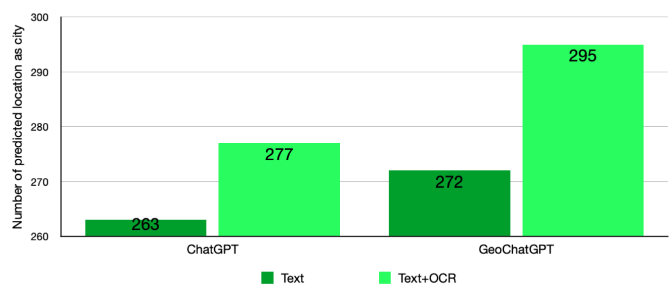

Figure 3: Admin level discrete distribution for city

4 EXPERIMENTAL RESULTS

This section presents the different configurations for the experiments and evaluates the geolocation improvement achieved by each setting. Specifically, we contrast the results when adding the extracted text from images using OCR into social media posts and having only text from social media posts using ChatGPT and GeoChatGPT.

	ChatGPT			**GeoChatGPT**		
	Text	Text+OCR	Delta	Text	Text+OCR	Delta
Precision	99.25	99.75	**0.5**	99.52	99.53	**0.01**
Recall	98.5	98.5	**0**	99.89	99.89	**0**
F1 score	98.75	99	**0.25**	99.71	99.71	**0**

Table 1: Delta improvement to quantify the enhancement when adding OCR to text at the event level

To evaluate our proposed approach, we measured the number of detected locations at different levels, including event, city, and neighborhood level. Initially, we focused on the evaluation at the event level. The event level refers to four locations of disasters in our dataset (Haiti earthquake, European floods, Croatia earthquake, and Catania floods). We used macro average for several performance metrics, including precision, recall and F1 score, to assess the delta improvement. Delta improvement is a quantitative measurement, including Precision, Recall, and F1 scores that demonstrate the effectiveness of the proposed approach. Experiment results demonstrated an increase in most of measurements when adding OCR to text. Also, there is an augmentation in performance from ChatGPT to GeoChatGPT. These results confirm that geo-guided ChatGPT model is more accurate than ChatGPT for detecting locations. The improvement of precision and F1 score vary from 0.01 to 0.5, while there was no improvement for recall for both models (see Table 1).

Administrative levels depict the granularity of a location within a governmental structure [4]. Our hypothesis stated that incorporating both image-derived texts and post texts will enhance ChatGPT's capability, leading to a more accurate representation of administrative boundaries. Specifically, by utilizing text extracted from images, we anticipate an improved capability to geolocate images down to the street and building number level.

[3] https://aws.amazon.com/id/rekognition

[4] https://wiki.openstreetmap.org/wiki/Key:admin_level

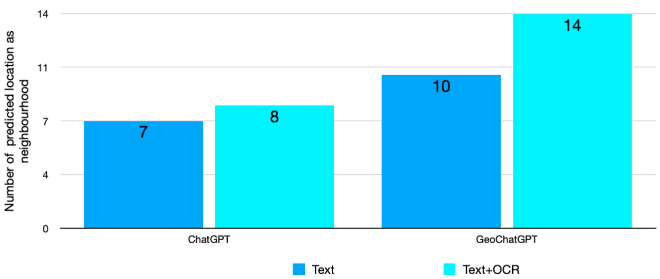

Figure 4: Admin level discrete distribution for neighborhood

To determine if the improvement happened at the city level, we contrasted the ChatGPT and GeoChatGPT. The performance of a GPT model is shaped by the prompts it is given. We created a prompt featuring a question and answer format, where the answer specifies the name of a location and its administrative hierarchy (such as country, city, or road). In Figure 3, the vertical axis represents the number of cities that were detected, while the horizontal axis shows the models. Based on the experiment, we could see the improvement in two different configurations. The highest number of detected cities was reached by GeoChatGPT with text and OCR. Using that configuration, it can recognize 295 cities' names. The lowest number was obtained by ChatGPT with text only (Figure 3). This result corroborates the improvement consistency at city level using our approach.

Furthermore, we performed the comparison to investigate the reliability of our proposed approach at the neighborhood level. The validation process for this level was done manually since the dataset only covered the city level for the ground truth. The validation was conducted by finding the location based on its latitude and longitude using Google Maps [5] and comparing the location with the image attached to the social media post. The experimental results showed that the proposed approach can improve the number of detected locations. For ChatGPT, the improvement was increased from 7 locations to 8 locations. For GeoChatGPT, the improvement was raised from 10 locations to 14 locations. The highest result was obtained by GeoChatGPT with the combination of text and OCR. This model could detect 14 locations at the neighborhood level (Figure 4). These results indicate that the combination of text and OCR consistently improves the detected location quantity and precision using ChatGPT and GeoChatGPT for all administrative level (event, city, and neighborhood).

At the neighborhood level, our approach can detect a hospital name called "Ospedale Garibaldi" based on information from the image (see Figure 5) which proved useful for the geolocation. Meanwhile, in the text we have the following post "#Catania piove sempre sul bagnato #Sicilia" which does not contain a precise location. This example demonstrates that the relevant text in an image could be beneficial to know the location of a social media post.

To visualize the results, we created a disaster map. In the disaster map, each location has its own latitude and longitude. We drew several blue bounding boxes that represent the impacted area of disaster. The exact coordinates of the impacted area were obtained

[5]https://www.google.com/maps

Figure 5: Sample of the detected location at the neighborhood level (left:original image, right:detail point of interest)

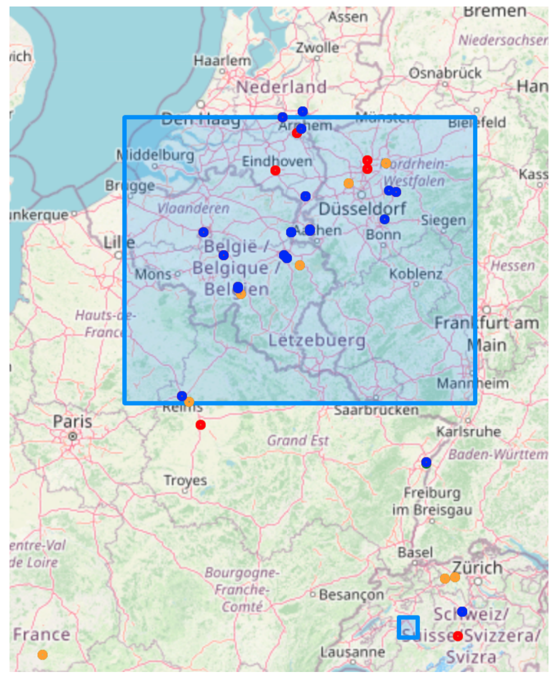

Figure 6: Disaster map created using social media data. Blue dots represent ChatGPT(Text). Red dots indicate ChatGPT(Text+OCR). Orange dots symbolize GeoChatGPT (Text). Green dots correspond to GeoChatGPT (Text+OCR)

from Copernicus Emergency Management Service (EMS) activation [6]. Based on the visualization, we observed that most of the predicted locations using our approach were within the bounding box (see Figure 6). These results also corroborate that a disaster map can be generated automatically using social media data. This disaster map is critically important for the first responders to understand the current situation in the impacted area of a disaster.

5 DISCUSSION

The findings from our study indicate that the integration of text extracted from images (OCR) in social media posts together with the post texts enhances overall geolocation performance. This improvement was observed consistently across both ChatGPT and

[6]https://emergency.copernicus.eu/

Figure 7: Location prediction from street signs (left:original image, right:detail point of interest)

GeoChatGPT models. From a practical perspective, teams responsible for operations, public information, and first responders could benefit from actionable insights derived from this research.

Based on our study, clear amenities (church, street signs, shops) are capable of showing specific locations, (Figure 7). Even though the detected location slightly deviates from the actual location, it could potentially be a starting point to predict a more precise location.

As a scientific contribution, this work provides evidences to practitioners regarding the use of artificial intelligence and social media integration for disaster response. Additionally, this approach could fill in the data quality and capability gaps in the theory of digital twin city for disaster management [6]. Furthermore, as an added practical contribution, we introduced a systems architecture for enhanced geolocation that can be utilized to augment other Large Language Models (LLM) as a data enrichment source. The effectiveness of this method was assessed by comparing the performance impact of incorporating text extraction from images into the social media posts' textual content.

While the outcomes of our experiment are encouraging, it is important to acknowledge the limitations of our approach. For instance, our approach provides several potential locations, but still need to select one precise location. To mitigate this issue, we suggest a hybrid approach that leverages both machine learning algorithms and human expertise (by specialists or crowdsourced) to disambiguate between potential locations. In terms of energy consumption, the use of an LLM model in the experiment also has an impact for the environment.

6 CONCLUSIONS AND FUTURE WORK

Social media users share critical information regarding the impacted area, people's condition, and broken infrastructure at the time of a disaster. However, this information is currently underutilized due to location limitations in social media posts.

In this work, we proposed a novel approach to predict location by combining text extracted from images in social media posts and posts' text. The combined text was subsequently input into ChatGPT to identify the locations mentioned within it. The text extracted from the image, especially the amenities, may contain useful information. The experimental results demonstrate that the proposed approach improves the quantity of detected locations and increases the level of precision.

As future work, we are interested in assessing how crowdsourcing could be integrated with this approach in order to disambiguate potential locations. Also, we aim to evaluate the prediction further, for instance, by measuring quantitatively the distance between the predicted location and the ground truth.

Acknowledgements

This work was funded by the Indonesia Endowment Fund For Education (LPDP) and CUI University of Geneva, Switzerland, and partially supported by the European Union's Horizon 2020 research and innovation programme K-HealthInAIR project with ID 101057693 and by the grant PID2022-136787NB-I00 funded by MCIN/AEI/10.13039/501100011033.

REFERENCES

[1] Firoj Alam, Muhammad Imran, and Ferda Ofli. 2017. Image4Act: Online Social Media Image Processing for Disaster Response. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (Sydney, Australia) *(ASONAM '17)*. Association for Computing Machinery, New York, NY, USA, 601–604. https://doi.org/10.1145/3110025.3110164

[2] Firoj Alam, Ferda Ofli, and Muhammad Imran. 2018. CrisisMMD: Multimodal Twitter Datasets from Natural Disasters. *Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM)* (2018).

[3] Jaime Alonso-Lorenzo, Enrique Costa-Montenegro, and Milagros Fernández-Gavilanes. 2016. Language independent big-data system for the prediction of user location on Twitter. In *2016 IEEE International Conference on Big Data (Big Data)*. 2437–2446. https://doi.org/10.1109/BigData.2016.7840880

[4] Amna Asif, Shaheen Khatoon, Md Maruf Hasan, Majed A. Alshamari, Sherif Abdou, Khaled Mostafa Elsayed, and Mohsen Rashwan. 2021. Automatic analysis of social media images to identify disaster type and infer appropriate emergency response. *Journal of Big Data* 8, 1 (Jun 2021). https://doi.org/10.1186/s40537-021-00471-5

[5] Carlo Bono, Barbara Pernici, Mehmet Oguz Mulayim, Jose Luis Fernandez-Marquez, Amudha Ravi Shankar, and Edoardo Nemni. 2022. TriggerCit: Early Flood Alerting using Twitter and Geolocation - a comparison with alternative sources. In *19th ISCRAM Conference*.

[6] Chao Fan, Cheng Zhang, Alex Yahja, and Ali Mostafavi. 2021. Disaster City Digital Twin: A vision for integrating artificial and human intelligence for disaster management. *International Journal of Information Management* 56 (2021), 102049. https://doi.org/10.1016/j.ijinfomgt.2019.102049

[7] Hafiz Budi Firmansyah, Carlo Bono, Jose Luis Fernandez-Marquez, Jesus Cerquides, and Valerio Lorini. 2023. Improving Disaster Response by Combining Automated Text Information Extraction from Images and Text on Social Media. In *Frontiers in Artificial Intelligence and Applications*. 320–329. https://doi.org/10.3233/FAIA230701

[8] Hafiz Budi Firmansyah, Jose Luis Fernandez-Marquez, Jesus Cerquides, Carlo Bono, Valerio Lorini, and Barbara Pernici. 2023. Enhancing Disaster Response with Automated Text Information Extraction from Social Media Images. In *IEEE Big Data Service*. https://doi.org/10.1109/BigDataService58306.2023.00017

[9] Hafiz Budi Firmansyah, Jose Luis Fernandez-Marquez, Jesus Cerquides, and Giovanna Di Marzo Serugendo. 2023. Single or ensemble model ? A study on social media images classification in disaster response. In *Proceedings of the 10th Multidisciplinary International Social Networks Conference* (Phuket,Thailand) *(MISNC '23)*. Association for Computing Machinery, New York, NY, USA, 48–54. https://doi.org/10.1145/3624875.3624884

[10] Chiara Francalanci, Paolo Guglielmino, Matteo Montalcini, Gabriele Scalia, and Barbara Pernici. 2017. IMEXT: A method and system to extract geolocated images from tweets — analysis of a case study. *2017 11th International Conference on Research Challenges in Information Science (RCIS)* (2017). https://doi.org/10.1109/rcis.2017.7956563

[11] Zixuan Liu and Xianwen Wu. 2023. Structural analysis of the evolution mechanism of online public opinion and its development stages based on machine learning and social network analysis. *International Journal of Computational Intelligence Systems* 16, 1 (Jun 2023). https://doi.org/10.1007/s44196-023-00277-8

[12] Valerio Lorini, Emanuele Panizio, and Castillo Carlos. 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 ICWSM. https://doi.org/10.36190/2022.1

[13] Valerio Lorini, Paola Rufolo, and Carlos Castillo. 2022. Venice Was Flooding ... One Tweet at a Time. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 382 (nov 2022), 16 pages. https://doi.org/10.1145/3555107

[14] Dat T. Nguyen, Ferda Ofli, Muhammad Imran, and Prasenjit Mitra. 2017. Damage Assessment from Social Media Imagery Data During Disasters. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (Sydney, Australia) *(ASONAM '17)*. Association for Computing Machinery, New York, NY, USA, 569–576. https://doi.org/10.1145/3110025.3110109

[15] Amudha Ravi Shankar, Jose Luis Fernandez-Marquez, Barbara Pernici, Gabriele Scalia, Rossy Mondardini, and Giovanna Di Marzo Serugendo. 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 (2019), 331–340. https://doi.org/10.5194/isprs-archives-xlii-3-w8-331-2019

[16] Desheng Wu and Yiwen Cui. 2018. Disaster early warning and damage assessment analysis using social media data and geo-location information. *Decision Support Systems* 111 (2018), 48–59. https://doi.org/10.1016/j.dss.2018.04.005

[17] Tan Yigitcanlar, Massimo Regona, Nayomi Kankanamge, Rashid Mehmood, Justin D'Costa, Samuel Lindsay, Scott Nelson, and Adiam Brhane. 2022. Detecting Natural Hazard-Related Disaster Impacts with Social Media Analytics: The Case of Australian States and Territories. *Sustainability* 14, 2 (2022). https://www.mdpi.com/2071-1050/14/2/810

[18] Chris Cundy Kristy Choi Ni Lao Wei Liu Gaurish Lakhanpal Ryan Zhenqi Zhou Yingjie Hu, Gengchen Mai and Kenneth Joseph. 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 (2023), 2289–2318. https://doi.org/10.1080/13658816.2023.2266495

[19] Lei Zou, Danqing Liao, Nina S.N. Lam, Michelle A. Meyer, Nasir G. Gharaibeh, Heng Cai, Bing Zhou, and Dongying Li. 2023. Social media for emergency rescue: An analysis of rescue requests on Twitter during Hurricane Harvey. *International Journal of Disaster Risk Reduction* 85 (2023), 103513. https://doi.org/10.1016/j.ijdrr.2022.103513