

The Geography of Conflicts: A Comparative Assessment of Web Data

Jan Bongard
Institute of Data Science
German Aerospace Center
Jena, Germany
jan.bongard@dlr.de

Xuke Hu
Institute of Data Science
German Aerospace Center
Jena, Germany
xuke.hu@dlr.de

Tobias Elßner
Institute of Data Science
German Aerospace Center
Jena, Germany
tobias.elssner@dlr.de

Jens Kersten
Institute of Data Science
German Aerospace Center
Jena, Germany
jens.kersten@dlr.de

ABSTRACT

The lack of trust is one of the most frequently addressed research gaps that hinders the practical deployment of social media information into decision-making process of disaster management. Furthermore, literature points out the lack of transferability of the developed methods to other countries and languages, the lack of long-term analyses, and the focus on specific online platforms.

In order to bridge the gap between research and practice, this paper highlights in addition the need for data fusion, which allows web and social media data to be combined with commonly used information structures.

Using the current Ukraine crisis as an example, this initial paper focuses on the fusion of social media, online news and humanitarian datasets. The goal is to evaluate the reliability of social media information. By using an aggregation methodology that structures the data based on spatio-temporal and semantic similarity, a binary reliability indicator is derived as a first prototype, which will serve as a basis for further developments. In addition, information quality indicators are presented that allow to evaluate the context created by the fusion and the results obtained from it.

Preliminary results reveal the beneficial effect of the obtained context and reliability indicator. Furthermore, the results show that social media make a significant contribution to the dissemination of unique information, both in terms of spatio-temporal and semantic aspects.

CCS CONCEPTS

• Information systems → Data extraction and integration.

KEYWORDS

Web Data Fusion, Reliability, Spatial Analysis, Data Quality

ACM Reference Format:

Jan Bongard, Tobias Elßner, Xuke Hu, and Jens Kersten. 2023. The Geography of Conflicts: A Comparative Assessment of Web Data. In *8th ACM SIGSPATIAL International Workshop on Security Response using GIS 2023 (EM-GIS '23)*, November 13, 2023, Hamburg, Germany. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3615884.3629424>

1 INTRODUCTION

The tremendous potential of web and social media data for disaster and incident management has been demonstrated in numerous studies for a wide variety of disaster types, emphasizing their real-time availability and informativeness, covering disaster mitigation, preparedness, response as well as recovery phases [13, 16, 17, 26]. In order to make the information easily accessible for practitioners, researchers provide demonstrators that showcase outcomes by using interactive dashboard with GIS components [2, 5, 14, 32].

Humanitarian organizations are incorporating this trend into their strategies, embracing their digital transformation, enhancing their capacity to leverage vast amount of data and to collaborate with the scientific community [21], but up to date, web and social media data is not systematically integrated into their decision-making process [30].

Recent literature reviews identify research gaps such as the insufficient integration of multiple online platforms, the lack of trust, the underrepresentation of countries that are prone and vulnerable to disasters, as well as lacking analysis in the field of long-term disaster recovery [7, 25]. Furthermore, [22] highlighted that in over 50% of the reviewed studies researchers did not find ways to validate their results, which completely ignores practical requirements for emergency response applications. Especially decision-making depends on the presence, awareness, and treatment of information about insufficient data quality [23]. Moreover, research on the fusion of web and social media data with common crisis information structures for a comprehensive situational picture is particularly deficient [9]. Specifically, the fusion of different data sources is useful to derive and ensure a high quality of information. Overall, it seems that data mining technologies are advancing more rapidly than the development of practical solutions for their implementation into emergency response decision management [34].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

EM-GIS '23, November 13, 2023, Hamburg, Germany

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0346-1/23/11.

<https://doi.org/10.1145/3615884.3629424>

Motivated by this, this study aims to investigate important aspects of the fusion of heterogeneous web data, by using aggregation techniques that are based on semantic similarity as well as on the geographic key concepts of space, time and scale. On this basis, a spatio-temporal and semantic reliability indicator is proposed. Using this technique, it is intended to identify to what extent heterogeneous web data can be utilized to improve the reliability of social media information in order to provide better and more complete insights for the decision making process.

The idea is that single social media posts can be seen as reliable, if they coincide temporally and spatially with reliable sources, and, moreover, if there is a high semantic similarity between them. Heterogeneous web data can be structured according to this concept and form a context from which reliability can be derived. In addition, information quality dimensions are presented that help to evaluate the quality of the fusion and its results. To obtain spatial information from a wide range of web and social media data, the study utilizes techniques of geographic information extraction from texts and geocoding.

This initial study is presenting preliminary results of the described fusion technique by using the current Ukraine crisis as an example. Web data from online news aggregators, Twitter, Telegram, as well as humanitarian datasets from ACLED and Bellingcat were compiled for the study period ranging from June to December 2022.

Under the aspect of geographically enriched web data, suitable information quality dimensions are presented in the following chapter. Subsequently, related work is discussed in Chapter 3 and common fusion techniques in the field of disaster response are presented. In chapter 4, the different data sources and their acquisition are presented in detail. Chapter 5 is dedicated to preprocessing and the proposed fusion technique creating and evaluating context and reliability. While Chapters 6 and 7 focus on the presentation and discussion of the results, Chapter 8 concludes the study.

2 INFORMATION QUALITY DIMENSIONS IN FUSION MODELS

Incorporating information quality into fusion models is a rather difficult task, since information quality can be expressed in various dimensions, and different taxonomies are available to categorize them [24, 33]. In addition, it is a challenge to combine the variety of quality dimensions for the representation in decision management, as the representation should remain understandable, complete and easy to interpret.

Following the detailed glossary provided by Gutiérrez et al [10], four important dimensions and requirements for the application of geographically enriched web data are extracted, interpreted and adapted for the present work:

- **Reliability:** Since information and its sources are not equally reliable, the main challenge for the integration of web and social media data is to enhance and to derive their trustworthiness and credibility to avoid decreasing performance in fusion and failures in decision making [23]. Reliability can be assessed through a comparison with a context that is created by the integration of complementary sources.

- **Geographic Accuracy:** Refers to the degree to which information matches objects or places. Geocoded information comes with different geographic resolutions for which a unifying reference or a multiresolutional approach needs to be found. In addition, geolocations derived from text may be incorrect.
- **Information Availability:** Given a specific location, information can be either available or not. Moreover, availability is time dependent, because unavailability does not imply that information will remain unavailable indefinitely.
- **Completeness:** Missing or partial information can be approached from two perspectives. From a geographic viewpoint, a source may focus on a particular region and overlook others. From a content perspective, a source may focus on certain topics, or present them only very superficially, so that semantic information is not available for fusion.

Besides those four aspects, a subjective relevance is particularly important for decision making, as it reflects the current value of the information for a specific user and is thus dependent on his information needs, tasks and experience. The role-dependent information need is mostly referred to as relevance. A focus on relevant information can improve the outcome of the fusion and lessens the computative effort for vast amounts of web data, but relevance is also a temporal attribute, because irrelevant information can become relevant later on or vice versa [23].

Depending on the user-defined relevance, a specific context comes into focus. This context is based on a composition of information sources which is influenced by their geographic accuracy, availability and completeness. Therefore, these dimensions have an direct impact on the creation of reliability. Consequently, it is necessary to identify the context for establishing reliability and to quantify the contribution of different information sources.

3 RELATED WORK: FROM THEORY TO PRACTICE

Besides image processing, decision support is one of the most common application areas for information fusion, especially with the goal of reducing uncertainty, and increasing reliability and accuracy [10]. However, particularly in the area of disaster response, studies related to information fusion consider only a limited number of different data sources and only few consider their quality dimensions. In the following, literature will be presented that uses aggregation techniques as well as statistical and probabilistic operators to generate or incorporate reliability.

[30] introduces a new conceptual approach that uses tweet coordinates and it's content to derive a four-step categorization scheme to identify and filter reliable and relevant messages for emergency response during a disaster scenario. The assignment to the categories was based on their location within a disaster declared district, their spatial proximity to reports from trained staff on site and on their content-related affiliation to a disaster relevant topic. While the presence of spatial proximity was determined using a hexagonal grid, a LDA topic modeling was chosen to aggregate and identify disaster related tweets. Validation was performed using satellite-based flood information, and found that 79% of tweets send from

disaster affected areas were classified as highly or moderately relevant and reliable. The result highlights that a small number of reliability certifiers (e.g. trustworthy reports from on-site) can be key in filtering even large amount of "reliable" tweets.

[27] presents a methodology to augment ground sensor data and SAR satellite imagery with non-authoritative data such as geotagged tweets, google photos, road closures or traffic cameras in order to fill information gaps during emergency. For each data source a layer is generated, using kernel density smoothing to spatially extend the representation of non-authoritative data. The different layers are then merged with a weighted sum overlay, where the weighting is based on a user-selected criterion, such as the reliability of the data source which in turn is based on factors like confidence in the producer, accuracy of the geolocation or sensibility of the method to detect affected areas. Finally, Kriging is used to create a geostatistical representation of the merged layers to derive a crisis estimation map.

[12] combined remote sensing data with geotagged tweets to enhance probability mapping. The proposed technique uses an inverse distance weighted height filter to create probability index distribution (PID) layers from digital elevation models based on the twitter coordinates. A Gaussian Kernel extracts moisture weights from satellite moisture products for each PID layer. Ultimately, a normalized flood probability map is generated by combining the moisture-weighted PID layers.

4 DATA

With the invasion of Ukrainian territory by Russian troops in February 2022, the previously regionally bounded armed conflict developed into a highly dynamic war. In particular, targeted attacks on civilians and civilian infrastructure as well as a high volume of refugees lead to a long-lasting complex humanitarian situation. For the period from June 1st to the end of December 2022, social media messages, news and events from various sources were collected for the study and are briefly introduced in the following:

GDELT & EMM. The Global Database of Events, Language, and Tone (GDELT) monitors and analyzes articles from around the world to provide insights into global events and trends. For the study period, entries of GDELT's knowledge graph were extracted from the English-language master file [31], if a direct mention of the Ukraine could be identified. The European Media Monitor (EMM), in contrast, is a system that explicitly analyzes sources from different European countries to provide insights into trends of European interest [29]. By utilizing the advanced search tool [18], information on news articles originating from Ukraine were obtained via web scraping. Based on the information provided by both services, the original online news websites could be visited and web-scraped. Duplicates were sorted out based on title and URL. A total of 186k news could be obtained.

X (Twitter). Tweets were retrieved through two different API endpoints using the researcher access. While the geo-search endpoint was used to obtain tweets originating from Ukraine, whereas the full-archive endpoint was used to acquire tweets based on crisis-related keywords (attachment) in Ukrainian. Around one million

tweets were obtained. In the following the two data sets are distinguished by the use of `tw_geo` and `tw_kwd`.

Telegram. Messages were scraped by building a public channel crawler that extracts and visits links to further channels. The starting point of the crawler was a seed of 15 channels, which were determined on the basis of `liveumap.com` and thus provide information on the course of the war. During the crawling process, channels in Ukrainian language were considered but no further prioritization was undertaken to steer the crawling process. 615 channels were extracted containing nearly 675k messages within the study period.

Bellingcat. The independent journalistic organization is known for its expertise in open-source intelligence (OSINT). Their platform [4] provides access to verified reports on incidents affecting civilians and civilian infrastructure in Ukraine providing event descriptions, exact coordinates of the incidents (within 150 m accuracy), and direct links to the information source. Since the majority of the information originates from social media, a total of 1224 messages from Twitter, Facebook and Telegram were acquired on this basis. Due to its characteristics, the dataset is particularly suitable for the validation of the geoparsing approach used.

ACLED. The Armed Conflict Location and Event Database is a global collection and crisis mapping project covering various types of conflicts, including battles, protests, riots, or violence against civilians. Individual events are characterized by the actors involved, source of information, subcategorization into several event types, fatalities, as well as time and coordinates. Using ACLED's export tool [1], around 24k events were received for the study period.

5 METHODS

5.1 Preprocessing

After their acquisition, only messages and news in Ukrainian, Russian or English were selected by a language detector for further processing. Subsequently, a Named Entity Recognition was performed using `spaCy` along with the corresponding language models [28]. Only texts which contain at least one location or geo-political entity were forwarded. Since news contain longer texts and therefore multiple place names, the dataset has been restructured so that each mention of a place name results in an entry that carries the corresponding sentence. Each distinct place name was then geocoded using the Nominatim API, an OpenStreetMap-based geocoder. From the potential candidates, the one with the highest importance score was selected and assigned to all entries belonging to the requested place name. Since `spaCy` and Nominatim are not suitable for the detection and disambiguation of fine-grained place names [11], only entries with a Nominatim place rank below 19 were used, such that spatial information up to the level of villages and suburbs was included, whereas levels below, such as streets and city blocks were excluded [20]. In order to structure the entries for GIS analysis, they were finally mapped to Ukrainian municipalities [8] by employing the bounding box received from Nominatim.

5.2 Reliability Indicator & Context Assessment

The aim is to assign social media messages with a binary reliability indicator. The approach followed here defines reliability as an

indicator that compares space, time and semantic similarity. Like reliability checks of a user, this approach is primarily concerned with trustworthiness rather than with checking specific claims. Reliability therefore does not pursue the claim of fact-checking.

Since individual information from Twitter and Telegram have little credibility, they are initially categorized as unreliable (0). News, being subject to a certain journalistic standard, are classified as reliable (1), together with verified ACLED and Bellingcat information. Reliability for social media is determined by verifying that information from reliable sources represents similar content in the same place at the same time.

In order to check semantic similarity a latent space representation is obtained for each information using a multilingual sentence embedding (LASER) [3, 6] that is forwarded to a non parametric Chinese Restaurant Process Clustering. For each incoming embedding the incremental clustering assigns the nearest cluster if their cosine similarity is above the pre-set threshold of 0.85. If no cluster is found in close proximity, a new cluster is spawned. Clusters are represented by the concept of central centroids [19], derived by averaging a fixed number of cluster representative embeddings that are most similar to each other and assuring robustness to outliers [15]. For each day and each municipality it is checked whether social media and reliable information are aggregated within a cluster, if so, the social media content is likewise classified as reliable.

Based on the aggregation, it is also possible to capture the pairings of the different sources within the clusters in order to identify sources with high semantic similarity over time, and determine the contribution of a reliable source to establish reliability for social media. Thus, for each reliable source, the number of social media messages that the respective source identifies as reliable is determined. This amount is considered in proportion to the total social media volume. In the following this proportion is described as "contribution to reliability" (CTR).

Moreover, municipalities can be identified on a daily basis that are addressed only by a single source. On the one hand, these geospatial outliers indicate that potentially user-relevant but not reliable content is available. On the other hand, they also identify potential information gaps of other sources. Hence, the contribution of each source to geospatial outliers should be determined here as well, called "contribution to outliers" (CTO) in the following.

Both measures can serve as indicators that provide meaningful insight regarding the quality dimensions completeness and information availability.

These relationships are examined in a twofold manner: First on the complete dataset, and second, on a dataset restricted to information containing the keyword "hospital" in the three respective languages. This is meant to simulate a potential user information need in the context of hospitals and to examine the resulting reliability and its context.

5.3 Validation

The place name extraction and geocoding task is evaluated by using the Bellingcat dataset. The dataset is suitable because of its precise annotated coordinates and its multilingual information provided by various social media sources. For each listed event, the corresponding messages were obtained from Telegram, Twitter and Facebook

Table 1: Source-dependent context characteristics (hospital use case)

Source	# Messages	# Reliable (%)	CTR	CTO
tw_geo	112	2 (1.78)	-	2.7
tw_kwd	247	1 (0.40)	-	9.62
telegram	961	155 (16.12)	-	32.06
emm	373	373	70.58	9.06
gdelt	2138	2138	23.52	43.51
bellingcat	10	10	0	0.31
acled	52	52	5.88	2.62

Table 2: Source-dependent context characteristics (entire data set)

Source	# Messages	# Reliable (%)	CTR	CTO
tw_geo	38,787	1106 (2.85)	-	6.12
tw_kwd	34,171	1075 (3.14)	-	6.67
telegram	275,939	87,634 (31.75)	-	61.90
emm	64,740	64,740	81.91	6.55
gdelt	240,830	240,830	17.82	2.90
bellingcat	307	307	0.02	0.03
acled	21,720	21,720	0.23	15.80

via webscraping. With the aid of a translation, messages containing a place name were manually identified. These messages serve as the basis to determine how many messages spaCy detects with a place name, how many place names can be geocoded by Nominatim, and ultimately in how many cases the bounding box received from Nominatim covers the precise Bellingcat coordinate. The validation of reliability and context creation via semantic similarity and clustering is done by means of a qualitative analysis. The focus of the evaluation lies on the classified reliability given the respective context that is to be investigated through the contribution of various sources (CTR) and the contribution to geographic outliers (CTO).

6 RESULTS

In this section, the results of the validation will be addressed, followed by the presentation of the derived reliability indicator, taking into account the context, as well as its spatial distribution over the entire study period.

The validation of the preprocessing indicated that spaCy was able to detect 86% of all existing place names in the Bellingcat dataset. Subsequently, 66% of all messages could be located using Nominatim. Thereby the bounding boxes obtained by Nominatim cover the Bellingcat coordinates in 78% of the cases.

The below translated daily cluster of the Kupyansk municipality is one example of the usability of the proposed spatio-temporal and semantic fusion approach to create a reliable and comprehensive context:

- telegram: "Kupyansk under missile attack: the occupiers targeted the hospital and the playground, Russian terrorists fired at the city with S-300 missiles, there were no casualties, the Ministry of Health reported. The children have not yet

gone out to play, and the hospital has stopped working due to threats."

- tw_kwd: "The day before yesterday, a 9-story building and a polyclinic with a school. Yesterday, a 5-story building and a hospital with a kindergarten. Russian inhumans are destroying the city of Kupiansk"
- emm: "Russian occupiers attacked a hospital and a playground in Kupiansk, Kharkiv Oblast, on 25 November."
- emm: "The hospital served the city itself and the Kupiansk district, taking care of about 120,000 people."
- emm: "Last year, Kupiansk hospital underwent major repairs, the Ministry said."

Table 1 provides information on the composition of the context and reliability for the hospital use case example. It is worth noting that GDELT and Telegram in particular provide much relevant information. However, most of the social media information from Twitter and Telegram cannot be verified as reliable. CTR indicates that humanitarian datasets contribute little to establishing reliability and a major portion can be verified by EMM, although gdelat contributes most of the information to the context. The CTO is based on the respective amount of information, so that sources that contribute much information also tend to cover municipalities as a single source.

Comparing these characteristics with the context of the entire data set in Table 2, it can be seen that in direct contrast to Twitter, a lot of Telegram information can be verified as reliable, especially through EMM. Furthermore, it is noticeable that the contribution to geographic outliers for GDELT is quite low compared to GDELT's contributed information volume.

Due to the strong connection between Telegram and EMM news reporting, a lot of reliable social media information can be found in embattled municipalities, depicted in Figure 1.

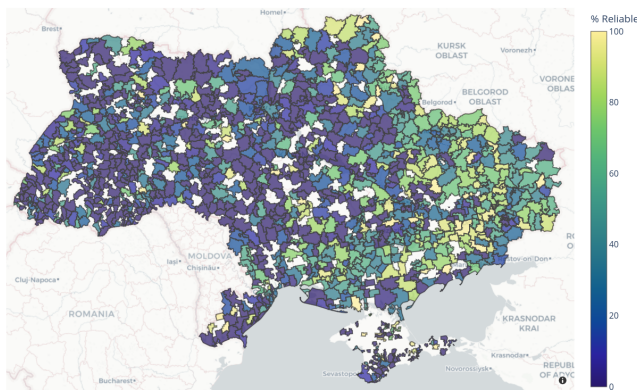


Figure 1: Percentage of reliable social media information in relation to the total information volume of social media posts (Twitter, Telegram) within Ukraine's municipalities throughout the study period Jun - Dec 2022

7 DISCUSSION

Web and social media data have an undeniable potential for disaster response management. Especially the combination of web scraping, geographic information extraction from texts and geocoding offers the possibility to include a variety of online platforms and to acquire data, independent of specific API services and their rate limits. This can help to analyze the impact of disasters in the long-term.

The validation of the preprocessing is in line with the results of other studies [11]. Although the combination of spaCy and Nominatim extracts and locates the majority of place names, there is still a need for further research for the transfer into practice. Tools and methods that perform better are often developed in English and do not meet the multilingual requirements. SpaCy is therefore used as a basic standard that allows the methods to be applied to the 25 other languages for which spaCy provides models.

An important part of the decision making process is the integration of meaningful information but also information and awareness about its quality. In particular, information fusion is a promising method to bring together individual sources in order to derive and ensure quality. Uncertainties and ambiguities can be reduced, which is particularly important in situations where decisions have to be made based on incomplete or contradictory information.

Up to date there is no systematic approach how different data sources of disaster response management can be included by fusion methods. Hence, this paper provides information quality dimensions specifically for the application of geographic enriched web and social media data focusing on reliability, geographic accuracy, information availability and completeness.

The results of the work show that by using spatio-temporal and semantic fusion social media information can be verified as reliable. The information quality is thereby increased and the goal of fusion is principally fulfilled. Nevertheless, the contribution to reliability indicate that primarily news from EMM tends to verify Telegram information. An unbalanced context reduces the usefulness of the fusion and ultimately the quality of the assigned reliability, since as many data sources with different characteristics as possible should ensure that uncertainties are avoided. This circumstance will be further investigated in future research.

Furthermore, it is evident that social media can contribute a high proportion of information that cannot be covered by news or humanitarian datasets either spatio-temporally or semantically. The example of the hospital use case also shows that although these geospatial outliers cannot be verified as reliable, they still have a relevance that corresponds to the user's information need. Compared to humanitarian datasets and news, social media has a high completeness (semantically and geographically) as well as information availability.

In future research, the goal is to integrate additional geographically widespread data sources like satellite data to assess, for example, the extent of civilian infrastructure destruction.

8 CONCLUSION

In conclusion, this research emphasizes the importance of integrating heterogeneous data sources to improve situational awareness and decision making in disaster management.

With the use of spatio-temporal and semantic fusion, a reliability indicator can be created. Thus, an additional measure for quality assessment is proposed which can ease the integration of unused data sources into decision management. The study also highlights the value of social media data resulting from its broad spatio-temporal distribution and semantic diversity. It could be shown that it is important to determine the influence of different sources on the context of the fusion. The influence depends on the user information needs, as well as on the spatio-temporal and semantic characteristics of the respective sources. The focus of further research is specifically to ensure that multiple sources are exploited to create a more balanced context that is less inclined to favor individual sources.

This preliminary study presents a novel approach that is capable of guiding future application-oriented research focusing on reliability and information quality of web and social media data.

REFERENCES

- [1] ACLED. [n. d.]. *Data Export Tool, The Armed Conflict Location and Event Database*. <https://acleddata.com/data-export-tool/>
- [2] Abeer AlAbdulaali, Amna Asif, Shaheen Khatoon, and Majed Alshamari. 2022. Designing multimodal interactive dashboard of disaster management systems. *Sensors* 22, 11 (2022), 4292.
- [3] Mikel Artetxe and Holger Schwenk. 2019. Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond. *Transactions of the Association for Computational Linguistics* 7 (2019), 597–610.
- [4] Bellingcat. [n. d.]. *Time Map: Civilian Harm in Ukraine*. <https://ukraine.bellingcat.com/>
- [5] Jan Bongard, Jens Kersten, and Friederike Klan. 2022. Searching and Structuring the Twitter Stream for Crisis Response: A flexible Concept to support Research and Practice. In *Proceedings of the 4th International Open Search Symposium (OSSYM)*.
- [6] facebookresearch. 2023. *LASER Language-Agnostic SEntence Representations*. <https://github.com/facebookresearch/LASER>
- [7] Muhammad Ashraf Fauzi. 2023. Social media in disaster management: review of the literature and future trends through bibliometric analysis. *Natural Hazards* (2023), 1–23.
- [8] OCHA Field Information Services Section (FISS). [n. d.]. *Ukraine - Subnational Administrative Boundaries via Humanitarian Data Exchange Portal*. <https://data.humdata.org/dataset/cod-ab-ukr>
- [9] Rob Grace. 2021. Overcoming barriers to social media use through multisensor integration in emergency management systems. *International Journal of Disaster Risk Reduction* 66 (2021), 102636.
- [10] Raúl Gutiérrez, Victor Rampérez, Horacio Paggi, Juan A Lara, and Javier Soriano. 2022. On the use of information fusion techniques to improve information quality: Taxonomy, opportunities and challenges. *Information Fusion* 78 (2022), 102–137.
- [11] Xuke Hu, Yeran Sun, Jens Kersten, Zhiyong Zhou, Friederike Klan, and Hongchao Fan. 2023. How can voting mechanisms improve the robustness and generalizability of toponym disambiguation? *International Journal of Applied Earth Observation and Geoinformation* 117 (2023), 103191.
- [12] Xiao Huang, Cuizhen Wang, and Zhenlong Li. 2018. A near real-time flood-mapping approach by integrating social media and post-event satellite imagery. *Annals of GIS* 24, 2 (2018), 113–123.
- [13] Mohammadsepehr Karimiziarani and Hamid Moradkhani. 2023. Social response and Disaster management: Insights from twitter data Assimilation on Hurricane Ian. *International Journal of Disaster Risk Reduction* 95 (2023), 103865.
- [14] Marc-André Kauffhold, Nicola Rupp, Christian Reuter, and Matthias Habdank. 2020. Mitigating information overload in social media during conflicts and crises: Design and evaluation of a cross-platform alerting system. *Behaviour & Information Technology* 39, 3 (2020), 319–342.
- [15] Jens Kersten, Jan Bongard, and Friederike Klan. 2021. Combining supervised and unsupervised learning to detect and semantically aggregate crisis-related twitter content. In *18th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2021*. 744–754.
- [16] Lingyao Li, Zihui Ma, and Tao Cao. 2021. Data-driven investigations of using social media to aid evacuations amid Western United States wildfire season. *Fire Safety Journal* 126 (2021), 103480.
- [17] Mahsa Moghadas, Alexander Fekete, Abbas Rajabifard, and Theo Köter. 2023. The wisdom of crowds for improved disaster resilience: a near-real-time analysis of crowdsourced social media data on the 2021 flood in Germany. *GeoJournal* (2023), 1–27.
- [18] European Media Monitor. [n. d.]. *EMM Newsbrief Advanced Search*. <https://emm.newsbrief.eu/NewsBrief/search/de/advanced.html>
- [19] Son Nguyen, Bao Ngo, Chau Vo, and Tru Cao. 2019. Hot topic detection on twitter data streams with incremental clustering using named entities and central centroids. In *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*. IEEE, 1–6.
- [20] Nominatim. [n. d.]. *Nominatim 4.3.0 Manual: Place Ranking in Nominatim*. <https://nominatim.org/release-docs/latest/customize/Ranking/>
- [21] International Federation of Red Cross and Red Crescent Societies. 2018. *Strategy 2030*. <https://www.ifrc.org/sites/default/files/2021-06/S2030-EN.pdf>
- [22] RI Ogie, S James, A Moore, T Dilworth, M Amirghasemi, and J Whittaker. 2022. Social media use in disaster recovery: A systematic literature review. *International Journal of Disaster Risk Reduction* 70 (2022), 102783.
- [23] Galina L Rogova. 2016. Information quality in information fusion and decision making with applications to crisis management. In *Fusion Methodologies in Crisis Management: Higher Level Fusion and Decision Making*. Springer, 65–86.
- [24] Galina L Rogova. 2019. Information quality in fusion-driven human-machine environments. *Information Quality in Information Fusion and Decision Making* (2019), 3–29.
- [25] Anita Saroj and Sukomal Pal. 2020. Use of social media in crisis management: A survey. *International Journal of Disaster Risk Reduction* 48 (2020), 101584.
- [26] Salim Sazzed. 2022. The Dynamics of Ukraine-Russian Conflict through the Lens of Demographically Diverse Twitter Data. In *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 6018–6024.
- [27] Emily Schnebele, Guido Cervone, Shamanth Kumar, and Nigel Waters. 2014. Real time estimation of the Calgary floods using limited remote sensing data. *Water* 6, 2 (2014), 381–398.
- [28] Spacy. 2023. *Trained Models Pipelines*. <https://spacy.io/models/uk>
- [29] Ralf Steinberger, Bruno Pouliquen, and Erik Van der Goot. 2013. An introduction to the europe media monitor family of applications. *arXiv preprint arXiv:1309.5290* (2013).
- [30] A Marcela Suarez and Keith C Clarke. 2022. A geographical and content-based approach to prioritize relevant and reliable tweets for emergency management. *Cartography and Geographic Information Science* 49, 5 (2022), 443–463.
- [31] Language The Global Database of Events and Tone. [n. d.]. *GDELT V2 Masterfile English*. <http://data.gdeltproject.org/gdeltv2/masterfilelist.txt>
- [32] Cagri Toraman, Izzet Emre Kucukkaya, Oguzhan Ozelik, and Umitcan Sahin. 2023. Tweets under the rubble: Detection of messages calling for help in earthquake disaster. *arXiv preprint arXiv:2302.13403* (2023).
- [33] Richard Y Wang and Diane M Strong. 1996. Beyond accuracy: What data quality means to data consumers. *Journal of management information systems* 12, 4 (1996), 5–33.
- [34] Camila E Young, Erica D Kuligowski, and Aashna Pradhan. 2020. *A review of social media use during disaster response and recovery phases*. US Department of Commerce, National Institute of Standards and Technology.

A APPENDIX

Twitter Full Archive API keyword query: Ambulance, Hospital, Pharmacy, Nurse, Doctor, Blood Donation, Train, Bus, Car, Kindergarten, School, College, University, Police, Fire, Fire Department, Radiator, Heating, Heater, Power, Generator, Fuel, Attack, Danger, Warning, Crime, Tank, Bomb, Shelling, Casualties

Received 15 September 2023; revised 19 October 2023; accepted 19 October 2023