

Social Media for Emergency Management: Opportunities and Challenges at the Intersection of Research and Practice

Valerio Lorini

European Commission, Joint Research Centre (JRC), Ispra, Italy
Universitat Pompeu Fabra, Barcelona, Spain
valerio.lorini@ec.europa.eu

Carlos Castillo

Universitat Pompeu Fabra, Barcelona, Spain
chato@acm.org

Steve Peterson

Community Emergency Response Team,
Montgomery County, Maryland, USA
stevepeterson2@gmail.com

Paola Rufolo

Fincons S.p.A., Milano, Italy
paola.rufolo@ext.ec.europa.eu

Hemant Purohit

George Mason University, Fairfax, USA
hpurohit@gmu.edu

Diego Pajarito

University of Warwick, Coventry, United Kingdom
diego.pajarito-grajales@warwick.ac.uk

João Porto de Albuquerque

University of Warwick, Coventry, United Kingdom
j.porto@warwick.ac.uk

Cody Buntain

New Jersey Institute of Technology, New Jersey, USA
cbuntain@njit.edu

ABSTRACT

This paper summarizes key opportunities and challenges identified during the workshop “Social Media for Disaster Risk Management: Researchers Meet Practitioners” which took place online in November 2020. It constitutes a work-in-progress towards identifying new directions for research and development of systems that can better serve the information needs of emergency managers.

Practitioners widely recognize the potential of accessing timely information from social media. Nevertheless, the discussion outlined some critical challenges for improving its adoption during crises. In particular, validating such information and integrating it with authoritative information and into more traditional information systems for emergency managers requires further work, and the negative impacts of misinformation and disinformation need to be prevented.

Keywords

Crisis informatics, workshop report.

INTRODUCTION

While social media has radically reshaped many industries, how to make this abundant source of information and online behavior more accessible and usable for crisis management has remained a subject of debate – as remarked upon in a recent retrospective on crisis informatics trends and technology (Reuter et al. 2018). Part of this debate stems from possible disconnects between practitioners’ needs and researchers’ agendas, as practitioners believe social media has value for their domains, but key concerns that have inhibited these benefits remain unaddressed. The E1 Unit operating within the “Space, Security and Migration” directorate of the Joint Research Center (JRC), has sought to bridge these two perspectives by organizing a workshop on “Social Media for Disaster Risk Management: Researchers Meet Practitioners”. The workshop consisted of five panel sessions conducted online on November 30th and December 1st, 2020, and attended by 70 people.¹ Grounding this workshop in a set of emergency-related case studies, replete with numerous datasets of various sources and modalities, the workshop encouraged practitioners to share their perspectives with researchers about key aspects in which social media may provide utility in these events. Researchers shared their ideas on key technical and scientific challenges. Following this cross-disciplinary exchange, this paper summarizes and highlights some of the main areas identified by practitioners, recognizes the perceived barriers within them, and describes the gaps between practitioner requirements and the research that need to be conducted to overcome these barriers. In making this summary available, we work towards establishing a common ground between practitioners and researchers for discussing methods for channeling social media information in the most fruitful way. This approach could ultimately move the field ahead by orienting the significant research effort in this space toward solving fundamental issues that have inhibited practitioners in their use of social media in crisis management.

To organize this summary, we first layout the benefits and values practitioners perceive in using social media for crisis management. We then describe the main concerns practitioners indicate that have possibly inhibited realization of these benefits and the research related to them. These issues are not new to researchers and to a large extent they have been mentioned elsewhere (e.g. Reuter et al. 2018). Key concerns we highlight include:

- preventing negative consequences from misinformation and disinformation;
- validating social media information; and
- presenting crisis-relevant information from social media in a useful manner.

PRACTITIONERS’ PERCEPTIONS OF THE VALUE OF SOCIAL MEDIA

Over the workshop, practitioners identified four areas in which they saw social media providing significant value for crisis management:

- Real-time and ongoing situational awareness;
- rapid insights in the immediate aftermath of disaster;
- integration with heterogeneous, multi-modal data; and
- flexible value across the crisis management cycle.

Real-Time and Ongoing Situational Awareness. Social media’s increasingly ubiquitous and global presence, particularly during disasters and emergencies, has the potential to empower decision-makers in crisis management by providing access to timely and relevant information. While, for instance, earth observation data may be delayed while satellites wait for necessary orbital positioning, or forecasts may be inaccurate in time and space due to lack of sufficient observational data, citizen-generated information in social media is available 24/7/365 in real-time, all over the globe. This data can supplement risk assessment improving timeliness and efficiency of satellite based emergency mapping (Dottori et al. 2017). Recent advances in computational approaches to the massive scales of data made available through social media provide an opportunity to improve situational awareness for management and response teams.

Rapid Insights in the Immediate Aftermath of Disaster. In general, as time passes the conversation of social media tends to become more saturated with noise (e.g., jokes, political blaming, general negative sentiment, etc.) causing it to depreciate in value for practitioners. In past studies, the volume of relevant or actionable information was often considered as an indicator of an event, as it was assumed in general, that there is a spike of crisis-related

¹<https://publications.jrc.ec.europa.eu/repository/>

messages around a particular location or topic. Recent studies (Lorini et al. 2019; de Bruijn et al. 2020) take a more conservative approach and integrate authoritative forecasts as well as incorporate crisis responders' feedback (Purohit and Peterson 2020) in defining the information filtering frameworks, which can better extract and organize actionable information to inform decision makers while overcoming noise and misinformation.

An interesting case study presented during the workshop (the explosions in Beirut from August 4th, 2020) highlighted how images and videos posted by witnesses early after a disaster, can be valuable sources when no other visual depictions of a disaster exist, for instance before satellite or aerial images can become available.

The Beirut explosion is of particular note as the first image reporting the event appeared only 1 minute after the official time of the explosions (18:07 local time). This post was followed after 13 minutes by a second one showing damages at 1.5 km distance, and a third one 19 minutes later, showing the first rescuers operating in the field. In a short time, more and more images and videos appeared, providing a good understanding of the situation in the surrounding of the blast location. The images published after the first posts related to the occurrence of the event, carry the highest potential for impact assessment in terms of magnitude and location. Indeed, the average probability of images of being classified as not-relevant to the event confirms the timeline of social media activity when users, during the first hours, are mostly posting visual reports (witnesses) while later, they are joined by messages of solidarity (Rizk et al. 2019; Rufolo et al. 2021)

According to practitioners, beyond contributing to situational awareness, social media could be helpful in detecting sub-events within a crisis and monitoring how a crisis unfolds. Other use-cases range from performing damage assessment (Brouwer et al. 2017), to gathering insights about the compliance with measures and recommendations issued by authorities, and feedback about the impacts of these actions.

Integration with Heterogeneous, Multi-Modal Data. As evidenced by, among others, the Beirut explosion in August 2020, the analysis of crisis-related images and videos – in addition to the standard text-analysis processes – can help humanitarian response organizations to improve decision-making and prioritize tasks. Annotated imagery and the ability to extract and unify semantic and visual features respectively from a social media post's text and images can facilitate detection and damage assessment of a crisis in new ways (Imran et al. 2020). For example, leveraging the growing popularity of visual media from disparate sources like Instagram, YouTube and TikTok, can support detecting and assessing severity and damage from natural crises like floods, fires, landslides, earthquakes; man-made crises like industrial accident, fallout from conflict; as well as severity assessment of the damage to infrastructure and the impact on the population.

While multi-modal data such as images and video are becoming more popular, so too are the needs to integrate these various kind of data into unified presentation layers. Crisis management could be improved, for example, through careful integration of this social data with other technology-based data including geospatial analytics and sensor technologies. It would require software systems designed for decision support that can handle heterogeneous data from an array of platforms for social networking, media sharing, and community-driven navigation, among others, with authoritative information from radars, satellites, sensors, and other sources. This would require technical advancement to enable the capability of representing and integrating these sources in a way that is relevant for decision models of the information systems to support emergency managers.

Flexible Value Across the Crisis Management Cycle. The value of social media information varies with each phase of the crisis management cycle. For example, during the preparedness phase, static social media messages from past crises can be studied to learn what type of relevant user-generated content could be anticipated in future events with similar crisis types. During the response phase, social media messages can be used to identify the extent of an event (Nguyen et al. 2017) or for immediate damage assessment. Social media could also be used as part of an iterative procedure for assessing efforts made by practitioners in the recovery phase. During the mitigation phase, agencies can perform corrective measures responding to and recovering from future crises based upon lessons learned from the public social media messages communicated during past crises. For example, if evacuation orders were reported by the public through social media as “confusing”, agencies can develop clearer evacuation communication strategies in anticipation of a similar crisis in the future (Ruin et al. 2020).

BARRIERS TO LEVERAGING SOCIAL MEDIA IN CRISIS MANAGEMENT.

Despite its perceived value, several key concerns have reduced trust in – and therefore adoption of – social media analysis and related technology for crisis management.

Information Overload and Uncertainty in Automated Filtering. Information overload is a critical concern when dealing with social media data. Practitioners accustomed to one-directional dissemination of information to citizens are now exposed to vast amounts of data originating from the public, which precedes formal communications and

exposes practitioners to overwhelming volumes of information. Of the many social media messages available during a crisis, only a very small portion of it are valuable for emergency management (Olteanu et al. 2015; Purohit, Castillo, et al. 2018; Mccreadie et al. 2019).

Beneficial messages have some actionable qualities, for instance, such as a clear location, as well as understandable and detailed information about a situation. Most social media messages do not fulfill this criterion. For instance, a large volume of COVID-19 tweets contain political discussions that are only tangentially valuable for crisis risk management. Therefore, it becomes essential to use reliable implementations of well-designed technologies to filter, prioritize, and organize relevant data from social media sources for decision makers. Further, these technologies need to perform in such a way that decision-making processes can be improved within a timeframe that is expected for each crisis management phase, especially during the time-sensitive response phase.

Technologies for processing social media messages, including artificial intelligence, machine learning and natural language processing (NLP) methods, have been deployed to try to filter out irrelevant messages from social media streams. These technologies are valuable but are limited in what they can achieve. For example, it is not always clear for a person whether a message is useful or not for someone, or whether it may be useful in the future or for whom; if these assessments are difficult for humans, automatic methods trained on these human judgement are likely to encounter difficulty. Moreover, to obtain accurate models, many algorithms need event-specific human-labelled data that may not be available in the early hours of an emergency or disaster. In addition, such filtering or categorization processes need to be clarified to overcome the concerns related to the explainability and interpretability of the automated process.

Different people and communities that must respond to disasters can be served differently by social media information, depending on its source, the type of disaster, and the timing of the information. In general, the less time passes from the moment in which a social media message is posted to the moment in which it arrives to an officer or responder, the better.

Misinformation and Its Consequences. Misinformation (unintentional) and disinformation (intentional) may have huge consequences with social media, including loss of life or property. There is a great differential in responsibility between users of social media platforms posting this information, and officials who may disseminate a wrong or misleading message. There are always questions of authenticity around messages posted by users of social media platforms. The public is aware of these questions, but at the same time, expects that social media channels are monitored by authorities. Sometimes the actual harm from misinformation or disinformation can be small or clearly avoidable, such as fabricated images showing sharks swimming in a flooded highway,² but in other cases the consequences can be large, such as false accusations to individuals.³ Although the misleading or malicious content can be a small fraction of messages, if not removed or somehow flagged, their impact can be larger; also the fact that “bad” content is a small fraction of the total does not mean that “good” information abounds. Eventually, social media has some capacity for self-correction, especially in the immediate aftermath of an event, but misinformation can be extremely persistent and mislead the public as time goes by. For instance, vaccine hesitation has been, to some extent, a persistent message partially amplified by social media. Making matters worse for practitioners, traditional media outlets sometimes share the misinformation and disinformation. When stories are not thoroughly fact-checked but instead are prematurely disseminated to the public, the consequences of amplifying false or misleading information may harm a crisis management agency’s response and recovery efforts. Time and resources need to be diverted from crisis response tasks to address false or misleading information.

Validation and Verification. Practitioners want to avoid the risk of credibility loss resulting from communicating false information. To avoid this, it is imperative they check information for accuracy prior to dissemination. Also, if practitioners are to make decisions based on social media messages, the selected content must be clear, accurate and trustworthy. Validating and verifying information prior to taking action also reduces the risk of inaccurate or sub-optimal allocation of resources. False alarms and misleading or outdated information are an everyday challenge for emergency management professionals. For instance, fire fighters might mobilize to attend alerts that end up being false alarms. Practitioners are trained to validate and verify the information they receive, and one main component of this validation is the consultation with a succession of entities, and the integration of independent sources of information. In this regard, technologies that can collect and integrate heterogeneous data from various sources would be extremely valuable, particularly if they can integrate social media as a massive and dynamic source, with authoritative data. The processing of social media could be substantially improved if developed workflows can provide both timely and validated information. Efficient workflows often involve collaboration between human and automated elements. In this collaboration, the automated part would require transparency and understandable mechanisms that make them more trustworthy and easy to use by their human counterparts. As photos and videos

²McKenzie Sadeghi: “Fact check: Photo of shark on a flooded highway is faked”. USA Today, August 2020.

³“Reddit apologises for online Boston ‘witch hunt’.” BBC News, April 2013.

become more prominent elements within messages, they can become more valuable as a means to aid in verifying crisis information. Recent work is making headway in this area, with new datasets of crisis-related images from social media are now available for researchers (crisisMMD⁴). As a consequence, automated techniques to detect old, edited, or fabricated images are as important as methods for validating and verifying textual content.

Formatting Information. A social media messages stream must be properly formatted before being provided to practitioners, to increase the potential for its use in decision-making. A key element of this formatting is the inclusion of geospatial information. Messages must be associated with places, regions, sites, or roads/routes/paths. In many cases, this association needs to offer high-accuracy and high-resolution, which is challenging due to the lack of an exact geographical reference in many social media messages. Locations mentions in the text of social media posts can be extracted as a place name candidates using part-of-speech taggers and searched against gazetteers (Halterman 2017). However, this approach depends on availability of data and gazetteers, and is often biased towards messages in English, limiting its applicability at a global scale. Experiences from humanitarian and collaborative mapping among other crowdsourced data collection initiatives may contribute to adding the geographical reference to social media. At the same time, more precise geolocation of messages may involve additional privacy issues when releasing fine-grained microdata. Practitioners may need data in various modalities. The first preference expressed is in a format that can be integrated within Geographical Information Systems (GIS) for visual display on a map. Secondly, in “raw” universal textual formats such as Comma-Separated Values (CSV) or simple text files. Thirdly, in other kinds of structured or tabular form. Practitioners may also need other kinds of data according to the type of crisis. Different crisis events may require different types of reporting from social media, data formats, and levels of summarization. This highlights a need for technologies that have aspects or features that depend on the type of disaster. **Multilingualism** Finally, in many cases, multi-lingual data is available and needs to be handled. There is, hence, also a requirement for methods that can process and collect data in multiple languages to create summaries that can be of use to the needs of different communities inhabiting a city or region. Multilinguistic tech solutions are very important because information posted in one language often is not available in another. However, practitioners agrees also that the pursuit of a monolingual solution that benefits the larger number of persons should not delay progress in the geographical area in which it was initially designed to be implemented. Fine tuning multilinguistic technologies such as embeddings or NLP multi-lingual pre-trained general purpose models to replicate/mirror the solution in other languages could satisfy this challenge.

TECHNICAL CHALLENGES AND FUTURE STEPS

Many technical challenges were named during the workshop, including the development of research, methods, and systems, to:

- Extract, transfer, and load heterogeneous data from various sources, particularly authoritative ones, and reliably integrating them into the real-time workflows, systems and tools used by practitioners for decision support.
- Automatically appraise the quality of an information source, or assist in the validation of a piece of information, be it a text message or an image or video.
- Recognizing and categorizing messages where human annotators disagree on the usefulness of the message, to generate a signal of ambiguity that can be further studied or used.
- Automatically place social media messages in time and space in an accurate manner: geocode them precisely and with high-resolution, and determine if they are timely or refer to some past or future event.
- Summarize social media messages authored in multiple languages.

As future steps, we plan to deepen our study of the gap between research and practice in this space by surveying a number of experiences of usage of social media during emergencies and then preparing, based on that outcome, a manual of best practices that can be useful for researchers and practitioners. During the workshop, it emerged how, as a fundamental activity, it would be beneficial to focus on shortening the distance between practitioners and researchers practical scopes. Nearly two different languages are spoken when it comes to disasters, and the two groups look at disasters through entirely different lenses. Thus, there needs to be more emphasis on a liaison role who can translate the terminology each one uses, describes the decision-making process and the actors involved in an emergency response.

⁴<https://crisisnlp.qcri.org/crisismmd>

Since the participants expressed their interest in keeping the discussion alive, a suggested further step is designing a strategy to lessens the amount of time invested by the practitioner community to collaborate with the research community, particularly during or immediately following a disaster. We think that well-planned initiatives structured over time could be the best path towards a common language for framing, addressing and tackling the issues presented in this work. We believe this could finally cover the last mile that still prevents social media from being used as any other sensor in a crisis room.

REFERENCES

- Brouwer, T., Eilander, D., Loenen, A. van, Booij, M. J., Wijnberg, K. M., Verkade, J. S., and Wagemaker, J. (2017). “Probabilistic flood extent estimates from social media flood observations”. In: *Natural Hazards and Earth System Sciences* 17.5, pp. 735–747.
- de Bruijn, J. A., de Moel, H., Weerts, A. H., de Ruiter, M. C., Basar, E., Eilander, D., and Aerts, J. C. (2020). “Improving the classification of flood tweets with contextual hydrological information in a multimodal neural network”. In: *Computers & Geosciences* 140, p. 104485.
- Dottori, F., Kalas, M., Lorini, V., Wania, A., Pappenberger, F., Salamon, P., Ramos, M. H., Cloke, H., and Castillo, C. (Apr. 2017). “Testing an innovative framework for flood forecasting, monitoring and mapping in Europe”. In: *EGU General Assembly Conference Abstracts*. EGU General Assembly Conference Abstracts, p. 13394.
- Halterman, A. (2017). “Mordecai: Full Text Geoparsing and Event Geocoding”. In: *The Journal of Open Source Software* 2.9.
- Imran, M., Ofli, F., Caragea, D., and Torralba, A. (2020). “Using AI and Social Media Multimodal Content for Disaster Response and Management: Opportunities, Challenges, and Future Directions”. In: *Information Processing & Management* 57.5, p. 102261.
- Lorini, V., Castillo, C., Dottori, F., Kalas, M., Nappo, D., and Salamon, P. (2019). “Integrating Social Media into a Pan-European Flood Awareness System: A Multilingual Approach”. In: *Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management, València, Spain, May 19-22, 2019*. ISCRAM Association.
- McCreadie, R., Buntain, C., and Soboroff, I. (2019). “TREC Incident Streams: Finding Actionable Information on Social Media”. In: *Proceedings of the 16th International Conference on Information Systems for Crisis Response And Management*. May, pp. 691–705.
- Nguyen, D. T., Ofli, F., Imran, M., and Mitra, P. (2017). “Damage Assessment from Social Media Imagery Data During Disasters”. In: *2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 569–576.
- Olteanu, A., Vieweg, S., and Castillo, C. (2015). “What to Expect When the Unexpected Happens: Social Media Communications Across Crises”. In: *In Proc. of 18th ACM Computer Supported Cooperative Work and Social Computing (CSCW'15)*. EPFL-CONF-203562.
- Purohit, H., Castillo, C., Imran, M., and Pandev, R. (2018). “Social-EOC: Serviceability model to rank social media requests for emergency operation centers”. In: *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018*, pp. 119–126.
- Purohit, H. and Peterson, S. (2020). “Social Media Mining for Disaster Management and Community Resilience”. In: *Big Data in Emergency Management: Exploitation Techniques for Social and Mobile Data*. Springer, pp. 93–107.
- Reuter, C., Backfried, G., Kaufhold, M.-A., and Spahr, F. (2018). “ISCRAM turns 15: A Trend Analysis of Social Media Papers 2004–2017”. In: *Proceedings of the 15th International ISCRAM Conference* May, pp. 1–14.
- Rizk, Y., Jomaa, H. S., Awad, M., and Castillo, C. (2019). “A Computationally Efficient Multi-Modal Classification Approach of Disaster-Related Twitter Images”. In: *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*. SAC '19. Limassol, Cyprus: Association for Computing Machinery, pp. 2050–2059.
- Rufolo, P., Muraro, D., and Lorini, V. (2021). *Social Media Image Analysis in the Immediate Aftermath of the 2020 Beirut Blast*. Luxembourg (Luxembourg).
- Ruin, I., Shabou, S., Chardonnel, S., Lutoff, C., and Anquetin, S. (2020). “4 - When Driving to Work Becomes Dangerous”. In: *Mobility in the Face of Extreme Hydrometeorological Events 2*. Ed. by C. Lutoff and S. Durand. ISTE, pp. 91–118.