# Processing Social Media Messages in Mass Emergency: Survey Summary

Muhammad Imran
Qatar Computing Research Institute (HBKU)
Doha, Qatar
mimran@hbku.edu.qa

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

Fernando Diaz
Spotify
New York, NY, USA
diazf@acm.org

Sarah Vieweg*
Qatar Computing Research Institute (HBKU)
Doha, Qatar
sarahvieweg@gmail.com

## ABSTRACT

Millions of people use social media to share information during disasters and mass emergencies. Information available on social media, particularly in the early hours of an event when few other sources are available, can be extremely valuable for emergency responders and decision makers, helping them gain situational awareness and plan relief efforts. Processing social media content to obtain such information involves solving multiple challenges, including parsing brief and informal messages, handling information overload, and prioritizing different types of information. These challenges can be mapped to information processing operations such as filtering, classifying, ranking, aggregating, extracting, and summarizing. This work highlights these challenges and presents state of the art computational techniques to deal with social media messages, focusing on their application to crisis scenarios.

## KEYWORDS

social media; disaster response; emergency management

**ACM Reference Format:**
Muhammad Imran, Carlos Castillo, Fernando Diaz, and Sarah Vieweg. 2018. Processing Social Media Messages in Mass Emergency: Survey Summary. In *WWW '18 Companion: The 2018 Web Conference Companion, April 23–27, 2018, Lyon, France.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3184558.3186242

## 1 INTRODUCTION

Sudden-onset emergencies such as natural or human-induced disasters bring uncertainties and an increasing need for time-critical information for formal response organizations, affected communities and other concerned populations [13, 46]. In particular, immediately after sudden-onset events, when information is most needed, information may be scarce.

The growing adaption of Information and Communication Technologies (ICT) and social networking platforms such as Twitter and Facebook has created numerous opportunities to disseminate and consume time-critical information in the form of images, videos and textual messages during natural disasters and emergencies [47, 49]. Social media postings are useful for a number of crisis response and management tasks such as: gaining insights into the situation as it unfolds, identifying urgent needs of the affected communities, and assessing the severity of damage [8]. Hence, many formal disaster response or emergency management organizations are interested in finding ways to quickly and easily locate and organize the information that is most useful to them [48].

However, the time-critical analysis of high-velocity, high-volume social media streams involves solving multiple challenges, including real-time parsing of brief and informal messages, handling information overload, determining information credibility, and prioritizing useful information. These challenges can be mapped to classical information processing tasks such as filtering, classifying, extracting, aggregating, ranking, visualizing, and summarizing information. For example, automatic classification techniques help reduce information overload by filtering out irrelevant messages, and summarization techniques help gain situational awareness by extracting important summaries from the relevant messages.

In addition to the textual content of social media, multimedia content (images and videos) can be valuable for disaster response and management [36]. For instance, images posted can be used to evaluate the severity of damage to critical infrastructure such as buildings, bridges, and roads [33]. Furthermore, in areas without Internet access, satellites and Unmanned Aerial Vehicles (UAV) can collect imagery data to increase situational awareness [34]. However, in contrast to the main advances in text processing (e.g., natural language processing) techniques, less has been done for multimedia content processing.

This paper summarizes an extended survey [18] covering state of the art methods for processing social media to support various disaster response and management operations. We further include new research questions and challenges for social media information processing in crisis response.

*The author now works at Facebook.


*WWW '18 Companion, April 23–27, 2018, Lyon, France*

ACM ISBN 978-1-4503-5640-4/18/04..
https://doi.org/10.1145/3184558.3186242

# 2 INFORMATION NEEDS, DATA CHARACTERISTICS, AND ACQUISITION

## 2.1 Information needs

The information needs of formal emergency response organizations varies with their roles and duties, as well as with characteristics of their specific context and its evolution over time [15, 16]. For instance, humanitarian and governmental emergency management organizations seek high-level information about a disaster, such as the scale of the event, the number of people living in the affected areas, and the overall economic impact. In contrast, organizations such as local police forces and firefighters benefit from information concerning individual emergency cases, such as urgent medical emergency reports, and the location of severely injured or trapped people. Roughly, these two types can be characterized as seeking to understand “the big picture” versus finding “actionable insights.”

## 2.2 Data characteristics and acquisition

The factors that can trigger an increase in social media communications can be endogenous or exogenous. The former includes spontaneous popularity increases of an information “meme” due to information cascades or contagion. The latter includes large-scale events such as disasters, emergencies, and mass convergence events. Data quality, in terms of readability, grammar, and sentence structure, vary significantly across social media platforms. For instance, messages in Twitter are often brief, informal, unstructured, and contain grammar and spelling mistakes. Preprocessing this data is an essential step before it can be used for any computation.

For automatic data collection, most large social media platforms provide Application Programming Interfaces (APIs). Generally, these APIs can be categorized into two types: search APIs and streaming APIs. Search APIs provide access to archived messages up to a certain limit. Streaming APIs allow subscribers to continuously consume data from a real-time data feed (again data limits maybe imposed). Different data collection strategies can be used to collect data, depending on the information need expressiveness allowed on these APIs. For example, a geographical rectangle (or bounding box) can be defined to acquire geo-tagged messages, a boolean query composed of keywords can be defined to acquire messages matching such query, or a specific set of users can be specified to collect everything those users post.

## 2.3 Event detection

Social media data processing systems often start with the automatic detection of an event. Crisis situations can be expected (i.e., predicted or forewarned) or unexpected. Expected crises include extreme weather events such as storms and tornadoes, where information is usually broadcast publicly before the event happens. Unexpected events include earthquakes and technological disasters such as industrial accidents, which can be anticipated only to a very coarse degree. The Topic Detection and Tracking (TDT) research community has developed techniques based on data from news media and other sources [3] including story segmentation, topic detection, new event detection [5, 40, 41], link detection, and topic tracking [11, 45]. A popular approach to detect new events in a data stream is to look for sudden increases in the frequency (“bursts”) of sets of keywords. However, this approach has some shortcomings. For instance, many popular hashtags (such as #followfriday) do not represent real-world events. Other approaches include using wavelet-based clustering of frequency signals, topic clustering with meta-data analysis, and domain-specific approaches using predetermined rules [12, 50].

## 2.4 Retrospective versus live data processing

Once an event is detected, its continuous tracking to obtain real-time insights is important for crisis responders. To be effective during emergencies, the data should be processed in real-time, and not retrospectively. Live data analysis (online processing) is usually performed over a current stream of data relevant for an event, often provided in real-time or with a short delay. The trade-off between retrospective and real-time data analysis is a matter of accuracy versus latency. However, in time-critical situations, the urgency with which the output of an analysis is required remains high, thus online algorithms using stream processing architectures can play a vital role [19, 28].

# 3 CLASSIFICATION, CLUSTERING, AND SUMMARIZATION

Social media communications during disasters are now so abundant that it is necessary to sift through hundreds of thousands, even millions of data points to find useful information. In recent years, several text and multimedia processing methods have been developed to efficiently process this overwhelming amount of information.

## 3.1 Information classification

A key task is to separate relevant and useful information from irrelevant and/or noisy content. One well-known approach to this task is supervised classification, which requires labeled examples (usually thousands of them). The selection of the set of categories to be used is mainly driven by two main factors: the data that is present in social media during crises and the information needs of users [17].

A classification task broadly includes different typologies of messages *by information provided* (e.g., affected or injured people, infrastructure damage, urgent needs of affected population) [19], or *by information source* (e.g., eyewitness accounts, official government sources, TV, radio) [35], or *by information credibility factors* (e.g., fake news, rumors, disinformation, misinformation) [9], or *by temporal aspects* (e.g., using different temporal phases of an event including pre, during, and post) [11], *by geographical locations* (e.g., based on particular geographical area or places near to the disaster zone), or *by factual, subjective, or emotional content.*

## 3.2 Information clustering

Clustering, which is an unsupervised machine-learning technique, covers a family of methods that seek to identify and explain important hidden patterns in unlabeled data. To process social media data during crises, clustering can help gather semantically similar messages that need to be processed/examined by humans, for instance by displaying multiple equivalent messages as a single item instead of multiple ones [6, 42]. Clustering approaches can be used to find anomalies in crisis data streams and to discover human annotation errors to improve the quality of supervised classification process [17, 20].

### 3.3 Information summarization

Categorizing messages reduces the amount of information that has to be processed by humans, but the remaining set of messages, even in single categories, can still be overwhelming for crisis responders. Information summarization techniques generate text-based summaries of an event as it unfolds. Text summarization systems are usually designed to surface core topics discussed in a set of relevant documents by locating key sentences. Other systems can produce a summary by abstracting or generating new sentences [4].

In the context of disaster data summarization, temporal information summarization or update summarization methods can infer the importance of different sentences using disaster-specific features, including geo-location and language models representing the way in which people write in social media about a disaster [24]. Text summarization systems can generate evolving summaries of a disaster over time [23]. Other approaches adapt a two-step approach in which first critical sub-events are identified and then summarized [43]

### 3.4 Information veracity

While social media can be a valuable resource for crisis response efforts, one of the barriers that hinders its use as a resource is the lack of trust on its contents [14]. Determining information credibility is difficult, especially during an ongoing emergency situation, as it involves dealing with misinformation (unintentional), disinformation (intentional), and rumors (unverifiable). Automatic techniques have been proposed to determine information credibility on social media, most such methods are based on superficial characteristics of the messages and of the history of the users who post them [9, 39].

### 3.5 Semantic enrichment

Semantic enrichment can be used to add layers of semantics (e.g., meta-data) to information so that algorithms can understand and interpret its meaning. Named entity recognition is a well-known semantic enrichment technique useful to identify named entities (person, places, and organization) from a give piece of information (e.g., a tweet). The identified named entities can then be linked to concepts (in many cases real-world concepts). After named-entity linking, messages that have been semantically enriched can be used to provide faceted search, a popular approach to interactively search through complex information spaces. Such semantic techniques have been used in social media processing systems e.g., *Twitris* [38].

Automatic processing systems interact with each other for a variety of reasons, allowing users to engage in information "triangulation." However, allowing systems to communicate information in a unified way is a challenging endeavor. An effective way to enable such heterogeneous communications is through machine-understandable ontologies. In crisis informatics domain ontologies such as HXL [1], MOAC [2] define and categorize different concepts to facilitate a common understanding.

[1]http://hxl.humanitarianresponse.info/
[2]http://observedchange.com/moac/ns/

### 3.6 Multimedia processing

Despite the wealth of research studies on the processing of social media textual content, comparatively less has been done on multimedia (images and videos) content. Images shared on social media can significantly help in various disaster response tasks [36]. Additionally, messages containing images/photos have been found to be more helpful for other tasks such as estimating the epicenter of an earthquake [25].

Other approaches combine features from textual and imagery content to, for example, classify visually relevant and irrelevant tweets [10, 51] or to generate event timelines using image-text summarization [51]. Visual summaries of trending topics can be generated using a multimodal-LDA (MMLDA) method [7], and image classification can be used to support a number of crisis response tasks [2] such as critical infrastructure damage assessment [33] and irrelevant image filtering [30].

Satellite imagery and aerial imagery captured via unmanned aerial vehicles (UAVs) is increasingly considered important and useful for disaster response. Aerial imagery can be captured and processed faster and at a lower cost, in comparison to satellite imagery data, and can potentially help gain situational awareness during the early hours of a disaster event. Hybrid systems combining human computation and machine learning techniques have been proposed for UAV imagery processing [34].

## 4 SYSTEMS FOR SOCIAL MEDIA MONITORING AND PROCESSING

Many systems for processing social media during disasters have been developed at various levels of maturity, ranging from proof-of-concept to production-level systems. For instance, *Twitris* [38] provides support for automatic classification of tweets and performs semantic enrichment. *Emergency Situation Awareness (ESA)* [37] performs event detection, text classification, online clustring, and geotagging. *SensePlace2* [26] filters and extracts geographical, temporal, and thematic information from tweets. *Artificial Intelligence for Disaster Response (AIDR)* [19] uses supervised machine learning techniques to categories tweets into humanitarian categories in real-time.

Common elements in these systems include lists and timelines to show recent important messages and groups containing semantically similar messages, maps for geotagged messages, graphs and charts to show visual summaries (e.g., proportion of messages), and time series graphs to represent volume of words, topics, or concepts over time.

## 5 CURRENT RESEARCH CHALLENGES AND FUTURE DIRECTIONS

In this section, we suggest some research directions based on the survey.

### 5.1 Domain adaptation and transfer learning

Obtaining human-labeled data to train automatic classifiers takes time; using data from past disasters may yield good results, but using data from the current disaster usually yields better results. Collecting recent training data rapidly is a challenge, and failing to do so may introduce unwanted inaccuracies or delays in the processing of important information. While some recent studies have demonstrated the utility of labeled data collected from past (potentially similar) disasters [22, 29, 32], robust machine learning techniques tested in different types of disasters (floods, earthquakes, hurricanes etc.) have yet to be developed. Methods such as *domain adaptation* and *transfer learning* can be employed to utilize data from past events, possibly in combination with unlabeled from the new event.

## 5.2 Online and active learning

In contrast to the batch learning approach in which machine learning models are trained using a batch of training examples, an online learning technique updates models upon receiving each new training instance (e.g., by receiving new labeled data from digital volunteers that are active immediately following a disaster [27]). Online learning helps models dynamically adapt to new changes and patterns in the data. However, two core issues in online learning techniques that use crowsourcing are (i) which tasks (e.g., tweets) to select for crowdsourcing? and (ii) when is the suitable time to create each crowdsourcing tasks?

Due to the abundance of duplicate messages, particularly on Twitter, asking for labels for messages that closely resemble each others can potentially waste crowdsourcing work. Moreover, presence of duplicate messages in the training and testing sets may lead to misleading measurements of accuracy. Possible solutions include, for example, having an aggressive de-duplication strategy and employing an *active learning* technique to help select messages that are most likely to increase a system's accuracy. Additionally, techniques to uncover new concepts and to track concept evolution in the data stream can detect underlying data changes and trigger model re-training processes.

## 5.3 Information credibility

Rumor detection, dealing with misinformation and disinformation cases, and determining information credibility can address a core question of many humanitarian organizations: how much should they trust information found on social media. Despite several recent attempts, information credibility issues remain largely unresolved. Determining the credibility of information on social media is a challenging task. Techniques that try to tackle this issue should consider different factors such as what is the evidence? (e.g., search for more eyewitnesses), is there any bias associated with the provided information?, and consider a variety of different sources.

## 5.4 Applications of deep learning

Traditional text classification techniques rely on manually engineered features like cue words and TF-IDF vectors, and do not perform well when data features change due to high variability in social media data. Deep Neural Networks (DNNs) use distributed condensed representation of words and learn higher level abstract features automatically. Contrary to sparse discrete representations, distributed representations generalize well and thus can play an important role learning high variability in crisis data. Convolutional Neural Network (CNN) are useful for the classification of tweets into binary (information vs. not-information) and multi-class (affected individuals, infrastructure damage, donations) categories [31]. A proposed variation of CNN with multilayer perceptron (MLP-CNN) has also shown to perform well in multi-class classification [29]. Although deep neural network techniques are promising for social media data classification, they require large amounts of training data. Recent attempts to make crisis-related labeled data available are hopeful [21, 35], but much more remains to be done.

## 5.5 From situational awareness to actionable insights

Gaining *situational awareness* is a core task of many humanitarian organizations to gain insights from the disaster area. However, in some cases it is *actionable knowledge* in the form of pieces of information that are sought. These may include implicit and explicit requests related to emergency needs that should be fulfilled or serviced as soon as possible. Despite extensive research that mainly focuses on understanding and extracting updates contributing situational awareness from social media platforms, limited work has been done that focuses on understanding how actionable each piece of information is in a given context and for a given user. Computational techniques that can automatically identify actionable messages from a live data stream during emergency events, and that are able to assess their urgency and categorize them to match different information needs of humanitarian organizations, are essential for crisis responders to launch rapid relief efforts.

## 5.6 Humanitarian crises and health

Natural disasters, long-term wars, and conflicts create severe health issues to the people who live under these circumstances. According to the World Health Organization (WHO), over 130 million people around the world are suffering with critical health issues due to humanitarian crises.[3] Although this work does not extensively discuss the role of social media and computational methods to deal with humanitarian health issues, there exist several studies examining social media uses and the advantages of different computational techniques in this space. Al-garadi et al. presents a detailed survey on the use of social media to track pandemics that covers a number of machine learning techniques to process social media data. In a recent work [44], Twitter data has been used during different epidemics to find information that is potentially useful for various health organizations.

# 6 CONCLUSIONS AND TRENDS

Social media platforms add new functionalities and new users every day; at the same time, some communications are moving to platforms that are oriented to one-on-one conversations or small groups instead of broadcasting. Cameras to capture 360-degree and Virtual Reality (VR) video, as well as UAVs with cameras, are becoming more affordable and easy to use, expanding the user base that produces these contents. In the light of these trends, we could envision that different types of information and more complex forms of information will emerge, creating new challenges. We expect active research and new developments in this space using machine intelligence, human intelligence, or a combination of both.

[3]http://www.who.int/hac/donorinfo/highlights/highlights37/en/

## REFERENCES


[1] Mohammed Ali Al-garadi, Muhammad Sadiq Khan, Kasturi Dewi Varathan, Ghulam Mujtaba, and Abdelkodose M Al-Kabsi. 2016. Using online social networks to track a pandemic: A systematic review. *Journal of biomedical informatics* 62 (2016), 1–11.

[2] Firoj Alam, Muhammad Imran, and Ferda Ofli. 2017. Image4Act: Online Social Media Image Processing for Disaster Response.. In *Proc. of ASONAM.* 1–4.

[3] James Allan. 2012. *Topic detection and tracking: event-based information organization.* Vol. 12. Springer Science & Business Media.

[4] Javed Aslam, Fernando Diaz, Matthew Ekstrand-Abueg, Richard McCreadie, Virgil Pavlu, and Tetsuya Sakai. 2015. *TREC 2014 temporal summarization track overview.* Technical Report. NIST.

[5] Marco Avvenuti, Stefano Cresci, Andrea Marchetti, Carlo Meletti, and Maurizio Tesconi. 2014. EARS (earthquake alert and report system): a real time decision support system for earthquake crisis management. In *Proc. of KDD.* ACM, 1749–1758.

[6] Michele Berlingerio, Francesco Calabrese, Giusy Di Lorenzo, Xiaowen Dong, Yiannis Gkoufas, and Dimitrios Mavroeidis. 2013. SaferCity: a system for detecting and analyzing incidents from social media. In *Proc. of ICDM Workshops.* IEEE, 1077–1080.

[7] Jingwen Bian, Yang Yang, and Tat-Seng Chua. 2013. Multimedia summarization for trending topics in microblogs. In *Proc. of CIKM.* ACM, 1807–1812.

[8] Carlos Castillo. 2016. *Big crisis data: social media in disasters and time-critical situations.* Cambridge University Press.

[9] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information credibility on twitter. In *Proc. of WWW.* ACM, 675–684.

[10] Tao Chen, Dongyuan Lu, Min-Yen Kan, and Peng Cui. 2013. Understanding and classifying image tweets. In *Proc. of ACM MM.* ACM, 781–784.

[11] Soudip Roy Chowdhury, Muhammad Imran, Muhammad Rizwan Asghar, Sihem Amer-Yahia, and Carlos Castillo. 2013. Tweet4act: Using incident-specific profiles for classifying crisis-related messages.. In *Proc. of ISCRAM.*

[12] Courtney D Corley, Chase Dowling, Stuart J Rose, and Taylor McKenzie. 2013. Social sensor analytics: Measuring phenomenology at scale. In *Proc. of ISI.* IEEE, 61–66.

[13] John Harrald and Theresa Jefferson. 2007. Shared situational awareness in emergency management mitigation and response. In *Proc. of HICSS.* IEEE, 23–23.

[14] Amanda L Hughes and Leysia Palen. 2012. The evolving role of the public information officer: An examination of social media in emergency management. *J. of Homeland Security and Emergency Management* 9, 1 (2012).

[15] Amanda L Hughes, Steve Peterson, and Leysia Palen. 2014. Social media in emergency management. *Issues in Disaster Science and Management: A Critical Dialogue Between Scientists and Emergency Managers* (2014).

[16] Amanda L Hughes, Lise AA St Denis, Leysia Palen, and Kenneth M Anderson. 2014. Online public communications by police & fire services during the 2012 Hurricane Sandy. In *Proc. of SIGCHI.* ACM, 1505–1514.

[17] Muhammad Imran and Carlos Castillo. 2015. Towards a data-driven approach to identify crisis-related topics in social media streams. In *Proc. of WWW.* ACM, 1205–1210.

[18] Muhammad Imran, Carlos Castillo, Fernando Diaz, and Sarah Vieweg. 2015. Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)* 47, 4 (2015), 67.

[19] Muhammad Imran, Carlos Castillo, Ji Lucas, Patrick Meier, and Sarah Vieweg. 2014. AIDR: Artificial intelligence for disaster response. In *Proc. of WWW.* ACM, 159–162.

[20] Muhammad Imran, Sanjay Chawla, and Carlos Castillo. 2016. A robust framework for classifying evolving document streams in an expert-machine-crowd setting. In *Proc. of ICDM.* IEEE, 961–966.

[21] Muhammad Imran, Prasenjit Mitra, and Carlos Castillo. 2016. Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)* (23-28). European Language Resources Association (ELRA), Paris, France.

[22] Muhammad Imran, Prasenjit Mitra, and Jaideep Srivastava. 2016. Enabling Rapid Classification of Social Media Communications During Crises. *Int. J. of Inf. Sys. for Crisis Response and Management* 8, 3 (2016), 1–17.

[23] Chris Kedzie, Kathleen McKeown, and Fernando Diaz. 2014. Summarizing disasters over time. In *Proc. of KDD Workshops.*

[24] Chris Kedzie, Kathleen McKeown, and Fernando Diaz. 2015. Predicting Salient Updates for Disaster Summarization.. In *Proc. of ACL.*

[25] Yuan Liang, James Caverlee, and John Mander. 2013. Text vs. images: on the viability of social media to assess earthquake damage. In *Proc. of WWW.* ACM, 1003–1006.

[26] Alan M MacEachren, Anuj Jaiswal, Anthony C Robinson, Scott Pezanowski, Alexander Savelyev, Prasenjit Mitra, Xiao Zhang, and Justine Blanford. 2011. Senseplace2: Geotwitter analytics support for situational awareness. In *Proc. of VAST.* IEEE, 181–190.

[27] Patrick Meier. 2015. *Digital humanitarians: how big data is changing the face of humanitarian response.* CRC Press.

[28] Leonardo Neumeyer, Bruce Robbins, Anish Nair, and Anand Kesari. 2010. S4: Distributed stream computing platform. In *Proc. of ICDM Workshops.* IEEE, 170–177.

[29] Dat Tien Nguyen, Kamla Al-Mannai, Shafiq R Joty, Hassan Sajjad, Muhammad Imran, and Prasenjit Mitra. 2017. Robust Classification of Crisis-Related Data on Social Networks Using Convolutional Neural Networks.. In *Proc. of ICWSM.* 632–635.

[30] Dat Tien Nguyen, Firoj Alam, Ferda Ofli, and Muhammad Imran. 2017. Automatic Image Filtering on Social Networks Using Deep Learning and Perceptual Hashing During Crises. In *Proc. of ISCRAM.*

[31] Dat Tien Nguyen, Shafiq Joty, Muhammad Imran, Hassan Sajjad, and Prasenjit Mitra. 2016. Applications of Online Deep Learning for Crisis Response Using Social Media Information. *arXiv preprint arXiv:1610.01030* (2016).

[32] Dat Tien Nguyen, Kamela Ali Al Mannai, Shafiq Joty, Hassan Sajjad, Muhammad Imran, and Prasenjit Mitra. 2016. Rapid Classification of Crisis-Related Data on Social Networks using Convolutional Neural Networks. *arXiv preprint arXiv:1608.03902* (2016).

[33] Dat Tien Nguyen, Ferda Ofli, Muhammad Imran, and Prasenjit Mitra. 2017. Damage Assessment from Social Media Imagery Data During Disasters. In *Proc. of ASONAM.* 1–8.

[34] Ferda Ofli, Patrick Meier, Muhammad Imran, Carlos Castillo, Devis Tuia, Nicolas Rey, Julien Briant, Pauline Millet, Friedrich Reinhard, Matthew Parkan, et al. 2016. Combining human computing and machine learning to make sense of big (aerial) data for disaster response. *Big data* 4, 1 (2016), 47–59.

[35] Alexandra Olteanu, Carlos Castillo, Fernando Diaz, and Sarah Vieweg. 2014. CrisisLex: A Lexicon for Collecting and Filtering Microblogged Communications in Crises.. In *Proc. of ICWSM.*

[36] Robin Peters and Porto de Albuqerque Joao. 2015. Investigating images as indicators for relevant social media messages in disaster management. In *Proc. of ISCRAM.*

[37] Robert Power, Bella Robinson, John Colton, and Mark Cameron. 2014. Emergency situation awareness: Twitter case studies. In *Proc. of ISCRAM MED.* Springer, 218–231.

[38] Hemant Purohit and Amit P Sheth. 2013. Twitris v3: From Citizen Sensing to Analysis, Coordination and Action.. In *Proc. of ICWSM.*

[39] Vahed Qazvinian, Emily Rosengren, Dragomir R Radev, and Qiaozhu Mei. 2011. Rumor has it: Identifying misinformation in microblogs. In *Proc. of EMNLP.* Association for Computational Linguistics, 1589–1599.

[40] Alan Ritter, Oren Etzioni, Sam Clark, et al. 2012. Open domain event extraction from twitter. In *Proc. of KDD.* ACM, 1104–1112.

[41] Bella Robinson, Robert Power, and Mark Cameron. 2013. A sensitive twitter earthquake detector. In *Proc. of WWW.* ACM, 999–1002.

[42] Jakob Rogstadius, Maja Vukovic, CA Teixeira, Vassilis Kostakos, Evangelos Karapanos, and Jim Alain Laredo. 2013. CrisisTracker: Crowdsourced social media curation for disaster awareness. *IBM J. of Research and Development* 57, 5 (2013), 4–1.

[43] Koustav Rudra, Siddhartha Banerjee, Niloy Ganguly, Pawan Goyal, Muhammad Imran, and Prasenjit Mitra. 2016. Summarizing Situational and Topical Information During Crises. *arXiv preprint arXiv:1610.01561* (2016).

[44] Koustav Rudra, Ashish Sharma, Niloy Ganguly, and Muhammad Imran. 2017. Classifying Information from Microblogs during Epidemics. In *Proceedings of the 2017 International Conference on Digital Health.* ACM, 104–108.

[45] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2010. Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proc. of WWW.* ACM, 851–860.

[46] Diane H Sonnenwald and Linda G Pierce. 2000. Information behavior in dynamic group work contexts: interwoven situational awareness, dense social networks and contested collaboration in command and control. *Inf. Proc. & Mgmt.* 36, 3 (2000), 461–479.

[47] Kate Starbird, Leysia Palen, Amanda L Hughes, and Sarah Vieweg. 2010. Chatter on the red: what hazards threat reveals about the social life of microblogged information. In *Proc. of CSCW.* ACM, 241–250.

[48] Sarah Vieweg, Carlos Castillo, and Muhammad Imran. 2014. Integrating social media communications into the rapid assessment of sudden onset disasters. In *Proc. of SOCINFO.* Springer, 444–461.

[49] Sarah Vieweg, Amanda L Hughes, Kate Starbird, and Leysia Palen. 2010. Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *Proc. of SIGCHI.* ACM, 1079–1088.

[50] Jianshu Weng and Bu-Sung Lee. 2011. Event detection in twitter. *Proc. of ICWSM* 11 (2011), 401–408.

[51] Shize Xu, Liang Kong, and Yan Zhang. 2012. A picture paints a thousand words: a method of generating image-text timelines. In *Proc. of CIKM.* ACM, 2511–2514.