Crisis Sub-Events on Social Media: A Case Study of Wildfires

Shan Jiang [*1] **William Groves** [2] **Sam Anzaroot** [2] **Alejandro Jaimes** [2]

Abstract

Social media has been extensively used for crisis management. Recent work examines possible sub-events as a major crisis unfolds. In this paper, we first propose a framework to identify sub-events from tweets. Then, leveraging 4 California wildfires in 2018-2019 as a case study, we investigate how sub-events cascade based on existing hypotheses drawn from the disaster management literature, and find that most hypotheses are supported by social media, e.g., fire induces smoke, which causes air pollution, which later harms health and eventually affects the healthcare system. In addition, we discuss other unexpected sub-events that emerge from social media.

1. Introduction

Social media has become a powerful tool for emergency management during crisis events (Palen & Anderson, 2016; Imran et al., 2015). When a crisis occurs, critical information is generated on social media platforms from involved individuals sending messages to their families and friends (Huang et al., 2015; Dailey & Starbird, 2017; Simon et al., 2015). This information then gets rapidly spread due to the networked social structure of these platforms (Metaxa-Kakavouli et al., 2018; Kogan et al., 2015). During this process, this information is also monitored, filtered and utilized by emergency responders in order to take humanitarian actions and reduce the harmful effects of crisis events (Temnikova et al., 2015; Alam et al., 2018b; Imran et al., 2015).

However, a crisis is often not a stand-alone event. As a major crisis unfolds, a series of *sub-events* are likely to occur (Helbing et al., 2006; Berariu et al., 2015), e.g., a coastal earthquake can trigger tsunamis and cause building collapses, a hurricane can cause power outages and, later, delay emergency communications. Understanding these sub-events is crucial for crisis management, therefore is of interest in academic communities (Pohl et al., 2012; Abhik & Toshniwal, 2013; Xing et al., 2016; Srijith et al., 2017; Chen et al., 2018; Rudra et al., 2018; Meladianos et al., 2015; 2018; Bekoulis et al., 2019). Following previous work, we first propose a framework that couples a series of natural language processing methods to identify sub-events from tweets. Then, leveraging 4 California wildfires in 2018 as a case study, we apply this framework and investigate how sub-events cascade based on existing hypotheses drawn from the disaster management literature (Helbing et al., 2006), and find that most hypotheses of cascades are supported by messages on social media, e.g., fire induces smoke, which causes air pollution, which later harms health and eventually affects the healthcare system. In addition, we also discuss other unexpected sub-events, e.g., statements from different stakeholders, collateral climatic reports, etc. In sum, the contributions of this paper are two-fold:

- A framework to identify sub-events on social media.
- A case study of sub-events after California wildfires.

2. Related Work

A recent survey shows that nearly half of the population relies on social media during crisis to lookup or share information (Reuter et al., 2017). Accessing the right information (Zade et al., 2018) at the right time (Chauhan & Hughes, 2017) from the right person (Metaxa-Kakavouli et al., 2018) significantly impacts live-saving decision-making. Therefore, researchers have extensively investigated this issue from many aspects, e.g., algorithmically processing, filtering and classifying informative and actionable messages (Imran et al., 2014; Temnikova et al., 2015; Imran et al., 2015; Alam et al., 2018a; Nguyen et al., 2017; Zeng et al., 2016), designing interface and platforms that help with information accessibility and visibility (Leavitt & Robinson, 2017; Bica et al., 2017), integrating social media and contemporary technological infrastructures (Robinson et al., 2015; Dailey & Starbird, 2017; Soden & Lord, 2018), etc.

One specific focus of crisis management is in identifying sub-events as a major crisis unfolds (Imran et al., 2015). Earlier work clustered similar tweets together as sub-events but yielded less interpretable results (Pohl et al., 2012; Abhik & Toshniwal, 2013). Other unsupervised methods utilized

[*]Work done during an internship at Dataminr [1]Northeastern University, Boston, MA, USA [2]Dataminr, New York, NY, USA. Correspondence to: Shan Jiang <sjiang@ccs.neu.edu>.

Appearing at the International Conference on Machine Learning AI for Social Good Workshop, Long Beach, United States, 2019.

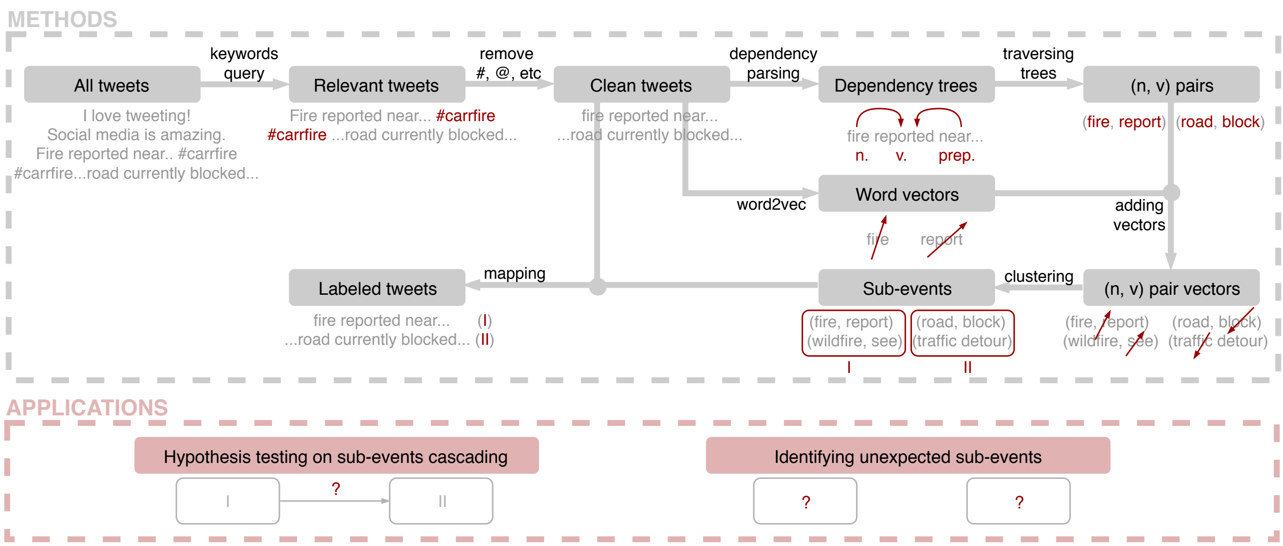


Figure 1. **Analytical framework.** Manually curated keywords are used to query the all tweets and retrieve relevant ones. After cleaning, we parse each tweet into a dependency tree and traverse it to get all associated (n, v) pairs. Meanwhile, we also vectorize words and (n, v) pairs in order to cluster similar (n, v) pairs together as a sub-event, and finally, map each tweet to its sub-events. These methods enable a set of applications, e.g., hypothesis testing on sub-events cascading and identifying unexpected sub-events.

topic models and deep learning methods (Xing et al., 2016; Srijith et al., 2017; Chen et al., 2018). A recent study found that sub-events can be framed as noun-verb pairs, e.g., building collapsed, help needed, etc. (Rudra et al., 2018), an observation we follow in this paper. In addition, there are also supervised methods for sub-events detection (Meladianos et al., 2015; 2018; Bekoulis et al., 2019), but these methods are validated in the football competition data context.

3. Framework

Our analytical framework, including methods for processing tweets and applications of sub-events, is shown in Figure 1.

3.1. Methods

We start by manually curating a small set of keyword queries to match the text of relevant Twitter messages (tweets), a method used by a number of previous studies (Temnikova et al., 2015; Olteanu et al., 2015; Alam et al., 2018a;b).[1] Then, we clean the text of tweets by removing Twitter handles (@), hashtags (#), URLs, etc.

A key observation made by previous research is that messages reporting sub-events consist two components: a noun (i.e., the entity that the sub-event is about) and a verb (i.e., the part that specifies what happened to the entity) (Rudra et al., 2018), e.g., fire reported, road blocked, etc. Therefore, we first identify these noun and verb pairs, i.e., (n, v) pairs, from clean tweets. Note that this identification is not a trivial task as associated nouns and verbs are not always directly adjacent to each other, e.g., in a tweet snippet "people need free, temporary accommodations" we need to extract "accommodations" and "need" as a (n, v) pair. To do this, we use a dependency parser implementation spaCy (Honnibal & Montani, 2017) to parse each tweet into a dependency tree using pre-trained models. Then, we traverse the tree and extract all child and parent nodes when one is a verb and the other is a noun.

The number of (n, v) pairs identified by this method is large, and we note that some (n, v) pairs represents very similar sub-events, e.g., fire reported, wildfire reported, fire see, etc. To cluster these (n, v) pairs together, we run word2vec on the tweet corpus (Mikolov et al., 2013), and average the vectors of the noun and the verb to represent the (n, v) pair.[2] Finally, we cluster similar (n, v) pairs together as a sub-event and map each tweet to its sub-events.

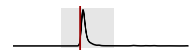
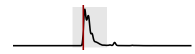
3.2. Applications

The above methods enable a set of applications, of which we explore two in this paper. The first application is hypothesis testing on sub-events cascading. Previous research on disaster management provides hypotheses on how sub-

[1]Note that this method is prone to false positives (i.e., retrieved irrelevant tweets) and negatives (i.e., unretrieved relevant tweets). We discuss this tradeoff during data collection in section 4.

[2]Several pre-trained models such as BERT (Devlin et al., 2018), spaCy (Honnibal & Montani, 2017), etc, are also tested to generate vectors but yield less satisfactory results. We find that for our task, word2vec implementation by gensim (Řehůřek & Sojka, 2010) yields similar results yet is much faster than fine-tuning BERT.

Table 1. **Wildfires and tweets.** 4 largest wildfires in California in 2018 are selected. Names of wildfires, curated queries, numbers and temporal distributions of tweets are shown. Distribution figures are fitted using kernel density estimates starting from Mar 2018 to Mar 2019. Wildfire starting dates (red lines) and selected time periods (shaded areas) are also shown.

Name	Query	Tweet	Distribution
Carr Fire	#carrfire OR ((#carr OR carr) AND (#fire OR fire OR #wildfire OR wildfire))	321K	
Mendocino Fire	#ranchfire OR #riverfire OR #mendocinocomplexfire OR ((#mendocinolakecomplex OR #mendocinocomplex) AND (#fire OR fire))	47K	
Camp Fire	#cafire OR #calfire OR ((#campfire OR #campfires OR #fire OR fire OR #wildfire OR wildfire) AND california)	1,014K	
Woolsey Fire	#woolseyfire OR #woolseyfires OR ((#woolsey OR woolsey) AND (#fire OR fire OR #wildfire OR wildfire))	580K	

events cascade after major crisis events, e.g., wildfire induces smoke, smoke causes air pollution, air pollution harms health, which later affects the healthcare system (Helbing et al., 2006). These hypotheses are reflected in social media and therefore can be tested by choosing certain measures, e.g., we can assume if one sub-event is caused by another, there should be an observable lag in temporal distributions of tweets regarding these sub-events.

The second application is in identifying unexpected sub-events. Given a set of known sub-events, e.g., hypotheses above, we can find unexpected sub-events by measuring the distance to known sub-events. Intuitively, a larger distance represents more semantic difference between the novel sub-event and the known sub-events, which could then be anticipated in the future, e.g., a wildfire could trigger statements from different stakeholders, collateral climatic reports, etc.

4. Case Study

We use wildfires as a case study. We selected the 4 largest wildfires in California in 2018, ranked by burned acres on Wikipedia (Wikipedia, 2018), including Carr Fire, Mendocino (Complex) Fire, Camp Fire and Woolsey Fire.[3]

We manually curate a text query for each wildfire and use them to retrieve all tweets that match these queries between Mar 2018 - Mar 2019, which covers the time duration of all selected events. The names of wildfires, curated queries, numbers and temporal distributions of tweets are shown in Table 1. Note that previous studies largely used relaxing sets of keywords which match a large set of tweets (Temnikova et al., 2015; Olteanu et al., 2015; Alam et al., 2018a;b), but also inevitably included a large potion of false positives. In contrast, we adopt a conservative approach by setting strict queries to ensure high relevance of retrieved tweets, each query is required to have either one or more event specific hashtags (e.g., #campfire), or contains a combination of keywords that specific event names and types (e.g., #camp AND #fire). We also filter out tweets that are measured to be far from the sample mean, by keeping only data within two standard deviations. This approach allow us to limit false positives to a negligibly small set: as shown in Table 1, the volumes of tweets are near 0 before each event and quickly reduce afterwards. However, this approach significantly increases our false negatives, i.e., many relevant Tweets do not follow the pattern we specified. Therefore, our data should be considered as a sample of explicitly relevant Tweets.

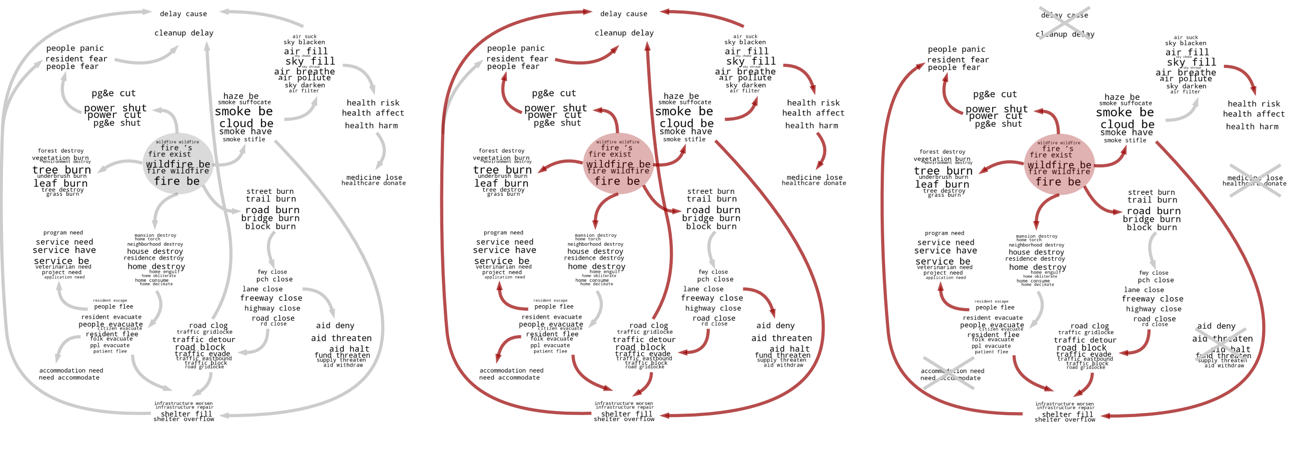
Ethics. We were careful to obey standard ethical practices during our data collection. All tweets used are publicly available and fully anonymized.

We then apply our framework on the dataset. The effectiveness of extracting (n, v) pairs is substantially evaluated by both domain experts and human annotators in (Rudra et al., 2018). And the effectiveness of clustering after word2vec is also evaluated by human, although from different domains (Fast et al., 2016; Jiang & Wilson, 2018). Despite that these evaluations do not necessarily guarantee the performance on our dataset, we focus on the results for the application part due to a lack of ground-truth.

4.1. Hypothesis testing on sub-events cascading

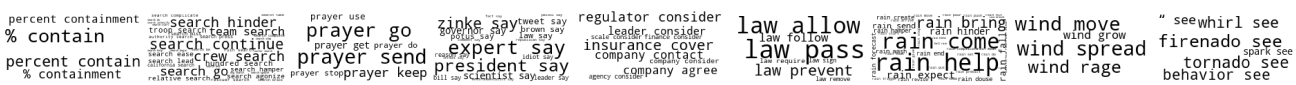
A hypothesized cascading network of sub-events is proposed in (Helbing et al., 2006), which describes a series of potential sub-events occurring after a typical wildfire. We reconstructed the network for our analytical purposes, as shown in Figure 2-(a). This network contains 18 nodes and 23 directed edges. First, we map each node to a sub-event by manually curating two (n ,v) pairs as seeds that describe each nodes, and then finding the clusters they belong to as

[3] Wildfires are chosen as a case study because its sub-events are well hypothesized in previous research in disaster management (Helbing et al., 2006). We select wildfires in 2018 to ensure our findings reflect the most recent emergency response environment. We choose California because it has a large user-base on Twitter and we choose largest fires because they are high-profile to emerge hashtags on Twitter.



(a) Reconstructed cascading network. (b) Evidence from Camp Fire. (c) Evidence from Carr Fire.

Figure 2. **Hypothesis testing on sub-events cascading.** Each cascade (edge) is tested by comparing the time lag between sub-events. The original cascading network and evidences from Camp Fire and Carr Fire are shown. A cascade is supported if both sub-events are present and the direction corresponds to the time lag. Supported cascades are highlighted with color.



(a) Contain. (b) Search. (c) Prayer. (d) Official. (e) Stakeholder. (f) Law. (g) Rain. (h) Wind. (i) Firenado.

Figure 3. **Identifying unexpected sub-events.** Examples of "unexpected", as measured by cosine similarity, sub-events are shown, include updates of wildfires on contained percentages and search efforts, prayers, statements from different stakeholders (e.g., officials, insurance, regulator, etc.), collateral climatic reports (e.g., rain, wind, firenado, etc.).

sub-events and filtering out irrelevant pairs. The resulting sub-events are visualized in Figure 2-(a), where the sizes of (n, v) pairs in the word cloud represent their frequencies in the corpus; Second, we test edges as cascades using time lags. The intuition is that if a sub-event occurs after another, there should be a lag between "when they are first tweeted". The time lags are measured as the time difference between the message posting times of the 1% temporal quantiles of the two sub-events. We report the results for Camp Fire and Carr Fire in Figure 2-(b-c) as they are two largest wildfires on different times. Due to space limitations, we omit results for Woolsey and Mendocino Fires.

Figure 2-(b) shows the evidence from Camp Fire. All 18 (100%) sub-events are identified and 20 of 23 (87%) cascades are supported. This includes complete cascading chains, e.g., fire induces smoke, which causes air pollution, which later harms health and eventually affects the healthcare system; fire induces power outages which lead to people panicking and eventually causes delay, etc. There are 3 cascades which are not supported, e.g., evacuation does not happen after home destruction, no significant lag between road burn and close, etc.

Figure 2-(c) shows the evidence from Carr Fire. 14 of 18 (78%) sub-events are identified and 13 of 16 (81%) cascades, minus the ones from or to unidentified sub-events, are supported. We observe a high degree of alignment between evidences from Carr Fire and Camp Fire for both supported and unsupported cascades, e.g., fire induces power issue and then panic, and evacuation does not happen after home destruction, etc.

4.2. Identifying unexpected sub-events

Besides the known sub-events in the hypothesized network, we also identify a number of "unexpected" sub-events. We measure "unexpectedness" by cosine similarities between vectors of (n, v) pairs. The intuition is to filter out sub-events that are related to the known. We keep all (n, v) pairs that are of distance more than 0.5 to all (n, v) pairs belonging to known sub-events and cluster them again.

Examples of these sub-events are shown in Figure 3. These sub-events include updates about wildfires including contained percentages and search efforts, prayers, statements from different stakeholders (e.g., officials, insurance, regulator, etc.), collateral climatic reports (e.g., rain, wind, firenado, etc.).

5. Conclusion

In this paper, we propose a framework to identify sub-events on social media during crises, and use the California wildfires during the 2018-2019 fire season as a case study for hypothesis testing on sub-events cascading and identifying unexpected sub-events.

Limitation. This paper describes an ongoing research project and has a few limitations, e.g., none of our methods are evaluated by domain experts, results vary strongly under different model selection and parameterization, pipeline methods accumulate errors at each step, etc.

Work in progress. We are currently working on evaluating our methods both qualitatively and quantitively using domain experts. And propose to extend these methods to understand sub-events of other crisis events, e.g., hurricanes, earthquakes, etc. In addition, we are trying to minimize human input and are making changes to systematically optimize model and parameter choices. Once we obtain a comprehensive labeled dataset, we also plan to use an end-to-end model to replace pipeline methods.

References

Abhik, D. and Toshniwal, D. Sub-event detection during natural hazards using features of social media data. In *Proceedings of the 22nd International Conference on World Wide Web (WWW)*, pp. 783–788. ACM, 2013.

Alam, F., Joty, S., and Imran, M. Graph based semi-supervised learning with convolution neural networks to classify crisis related tweets. In *Twelfth International AAAI Conference on Web and Social Media (ICWSM)*, 2018a.

Alam, F., Ofli, F., and Imran, M. Crisismmd: Multimodal twitter datasets from natural disasters. In *Twelfth International AAAI Conference on Web and Social Media (ICWSM)*, 2018b.

Bekoulis, G., Deleu, J., Demeester, T., and Develder, C. Sub-event detection from twitter streams as a sequence labeling problem. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 745–750, 2019.

Berariu, R., Fikar, C., Gronalt, M., and Hirsch, P. Understanding the impact of cascade effects of natural disasters on disaster relief operations. *International Journal of Disaster Risk Reduction*, 12:350–356, 2015.

Bica, M., Palen, L., and Bopp, C. Visual representations of disaster. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*, pp. 1262–1276. ACM, 2017.

Chauhan, A. and Hughes, A. L. Providing online crisis information: An analysis of official sources during the 2014 carlton complex wildfire. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI)*, pp. 3151–3162. ACM, 2017.

Chen, G., Xu, N., and Mao, W. An encoder-memory-decoder framework for sub-event detection in social media. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 1575–1578. ACM, 2018.

Dailey, D. and Starbird, K. Social media seamsters: Stitching platforms & audiences into local crisis infrastructure. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*, pp. 1277–1289. ACM, 2017.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

Fast, E., Chen, B., and Bernstein, M. S. Empath: Understanding topic signals in large-scale text. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI)*, pp. 4647–4657. ACM, 2016.

Helbing, D., Ammoser, H., and Kühnert, C. Disasters as extreme events and the importance of network interactions for disaster response management. In *Extreme events in nature and society*, pp. 319–348. Springer, 2006.

Honnibal, M. and Montani, I. spacy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. *To appear*, 2017.

Huang, Y. L., Starbird, K., Orand, M., Stanek, S. A., and Pedersen, H. T. Connected through crisis: Emotional proximity and the spread of misinformation online. In *Proceedings of the 18th ACM conference on computer supported cooperative work and Social Computing (CSCW)*, pp. 969–980. ACM, 2015.

Imran, M., Castillo, C., Lucas, J., Meier, P., and Vieweg, S. Aidr: Artificial intelligence for disaster response. In *Proceedings of the 23rd International Conference on World Wide Web (WWW)*, pp. 159–162. ACM, 2014.

Imran, M., Castillo, C., Diaz, F., and Vieweg, S. Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)*, 47(4):67, 2015.

Jiang, S. and Wilson, C. Linguistic signals under misinformation and fact-checking: Evidence from user comments on social media. *Proceedings of the ACM: Human-Computer Interaction*, 2(CSCW):82, 2018.

Kogan, M., Palen, L., and Anderson, K. M. Think local, retweet global: Retweeting by the geographically-vulnerable during hurricane sandy. In *Proceedings of the 18th ACM conference on computer supported cooperative work and Social Computing (CSCW)*, pp. 981–993. ACM, 2015.

Leavitt, A. and Robinson, J. J. The role of information visibility in network gatekeeping: Information aggregation on reddit during crisis events. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing (CSCW)*, pp. 1246–1261. ACM, 2017.

Meladianos, P., Nikolentzos, G., Rousseau, F., Stavrakas, Y., and Vazirgiannis, M. Degeneracy-based real-time sub-event detection in twitter stream. In *Ninth international AAAI conference on web and social media (ICWSM)*, 2015.

Meladianos, P., Xypolopoulos, C., Nikolentzos, G., and Vazirgiannis, M. An optimization approach for sub-event detection and summarization in twitter. In *European Conference on Information Retrieval (ECIR)*, pp. 481–493. Springer, 2018.

Metaxa-Kakavouli, D., Maas, P., and Aldrich, D. P. How social ties influence hurricane evacuation behavior. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):122, 2018.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

Nguyen, D. T., Al Mannai, K. A., Joty, S., Sajjad, H., Imran, M., and Mitra, P. Robust classification of crisis-related data on social networks using convolutional neural networks. In *Eleventh International AAAI Conference on Web and Social Media (ICWSM)*, 2017.

Olteanu, A., Vieweg, S., and Castillo, C. What to expect when the unexpected happens: Social media communications across crises. In *Proceedings of the 18th ACM conference on computer supported cooperative work and Social Computing (CSCW)*, pp. 994–1009. ACM, 2015.

Palen, L. and Anderson, K. M. Crisis informatics – new data for extraordinary times. *Science*, 353(6296):224–225, 2016.

Pohl, D., Bouchachia, A., and Hellwagner, H. Automatic sub-event detection in emergency management using social media. In *Proceedings of the 21st International Conference on World Wide Web (WWW)*, pp. 683–686. ACM, 2012.

Řehůřek, R. and Sojka, P. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pp. 45–50, Valletta, Malta, May 2010. ELRA. http://is.muni.cz/publication/884893/en.

Reuter, C., Kaufhold, M.-A., Spielhofer, T., and Hahne, A. S. Social media in emergencies: A representative study on citizens' perception in germany. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):90, 2017.

Robinson, J. J., Maddock, J., and Starbird, K. Examining the role of human and technical infrastructure during emergency response. In *Annual Conference for Information Systems for Crisis Response and Management (ISCRAM)*, 2015.

Rudra, K., Goyal, P., Ganguly, N., Mitra, P., and Imran, M. Identifying sub-events and summarizing disaster-related information from microblogs. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pp. 265–274. ACM, 2018.

Simon, T., Goldberg, A., and Adini, B. Socializing in emergencies – a review of the use of social media in emergency situations. *International Journal of Information Management*, 35(5):609–619, 2015.

Soden, R. and Lord, A. Mapping silences, reconfiguring loss: Practices of damage assessment & repair in post-earthquake nepal. *Proceedings of the ACM on Human-Computer Interaction*, 2 (CSCW):161, 2018.

Srijith, P., Hepple, M., Bontcheva, K., and Preotiuc-Pietro, D. Sub-story detection in twitter with hierarchical dirichlet processes. *Information Processing & Management*, 53(4):989–1003, 2017.

Temnikova, I. P., Castillo, C., and Vieweg, S. Emterms 1.0: A terminological resource for crisis tweets. In *Annual Conference for Information Systems for Crisis Response and Management (ISCRAM)*, 2015.

Wikipedia. 2018 california wildfires. https://en.wikipedia.org/wiki/2018_California_wildfires, 2018.

Xing, C., Wang, Y., Liu, J., Huang, Y., and Ma, W.-Y. Hashtag-based sub-event discovery using mutually generative lda in twitter. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

Zade, H., Shah, K., Rangarajan, V., Kshirsagar, P., Imran, M., and Starbird, K. From situational awareness to actionability: Towards improving the utility of social media data for crisis response. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):195, 2018.

Zeng, L., Starbird, K., and Spiro, E. S. # unconfirmed: Classifying rumor stance in crisis-related social media messages. In *Tenth International AAAI Conference on Web and Social Media (ICWSM)*, 2016.