Event Tracker: A Text Analytics Platform for Use During Disasters

Charles Thomas
2198970t@student.gla.ac.uk
University of Glasgow, UK

Richard McCreadie
richard.mccreadie@glasgow.ac.uk
University of Glasgow, UK

Iadh Ounis
iadh.ounis@glasgow.ac.uk
University of Glasgow, UK

ABSTRACT

Emergency management organisations currently rely on a wide range of disparate tools and technologies to support the monitoring and management of events during crisis situations. This has a number of disadvantages, in terms of training time for new staff members, reliance on external services, and a lack of integration (and hence poor transfer of information) between those services. On the other hand, Event Tracker is a new solution that aims to provide a unified view of an event, integrating information from emergency response officers, the public (via social media) and also volunteers from around the world. In particular, Event Tracker provides a series of novel functionalities to realise this unified view of the event, namely: real-time identification of critical information, automatic grouping of content by the information needs of response officers, as well as real-time volunteers management and communication. This is supported by an efficient and scalable back-end infrastructure designed to ingest and process high-volumes of real-time streaming data with low latency.

CCS CONCEPTS

• **Information systems → Decision support systems**.

KEYWORDS

Crisis Management, Social Media, Real-time Analytics

ACM Reference Format:
Charles Thomas, Richard McCreadie, and Iadh Ounis. 2019. Event Tracker: A Text Analytics Platform for Use During Disasters. In *Proceedings of the 42nd International ACM SIGIR Conference onResearch and Development in Information Retrieval (SIGIR '19), July 21–25,2019, Paris, France.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3331184.3331406

1 INTRODUCTION

A 68% rise in the use of social networking sites since 2005 [15] has introduced an abundance of real-time information online. This has enabled new ways for the public to contact response agencies [4], and grants those in the Emergency Management sector a new means of accessing potentially life-saving information. With a three-fold increase in natural disasters over the past 35 years [11], it is extremely important that emergency response agencies have the tools available to ensure that they can monitor social media streams in real-time and provide assistance to the public in a quick and effective manner.

Data posted on social media platforms regarding a disaster or world events could potentially provide a wide range of valuable information to emergency response services. For example, responding to cries for help, accessing first-hand observations or simply gaining an insight into public opinions could provide an added-value to these agencies. However, with the wealth of information now available, it is critical that the data can be easily visualised and navigated, ensuring that emergency services can quickly find and act upon key information.

There also lies potential in the connection of volunteers with formal emergency response officers during crisis situations. Systems designed to assist emergencies services during crises face typically a number of challenges, including a lack of coordination and communication between the officers and other formal entities as well as an unwillingness to engage and form relationships with digital volunteer groups [19]. A successful connection with both physical and digital volunteers could indeed aid response efforts immensely, with a wealth of potential skills and resources becoming available [20].

A range of systems have been developed in the past with the goal to support emergency response efforts during disasters [4]. These systems, however, have had little impact on the sector. It has been found that reasons behind this include the insufficient training of personnel, which require time to make adequate use of the platforms, the lack of guidelines for their use, and apprehension over social media trustworthiness [9].

In this paper, we propose a new system, Event Tracker, which aims to support the monitoring and management of events during disasters and crisis situations. The system intends to support three primary tiers of users - emergency response officers and their teams; physical and digital volunteers; as well as the victims during the crisis. Indeed, we argue that emergency response agencies should be given the tools to navigate large volumes of social media data with ease, supported by functionalities such actionable information feeds and criticality estimation. Moreover, volunteers should be able to provide information directly to these response agencies, either using their first-person knowledge of the situation, or aiding the data navigation and highlighting specific information and relevant details. Finally, victims during a disaster should be able to straightforwardly access advice and information from emergency services, ensuring that they are equipped with up-to-date information that could lead to their safety.

2 RELATED WORK

To-date, there have been a range of techniques and systems proposed to support crisis management. Below we provide a brief overview of key technologies and initiatives that are relevant to our proposed Event Tracker system.

SIGIR '19, July 21–25,2019, Paris, France

ACM ISBN 978-1-4503-6172-9/19/07…$15.00
https://doi.org/10.1145/3331184.3331406

Finding Relevant Information: A core function of any crisis management solution is getting key information to the right people. This is operationalised either via volunteer efforts [6] or through automatic technologies [18] for categorising and prioritising reports. Indeed, a survey [5] of approaches identified eight dimensions of categorisation, namely: by information provided/contained; fact vs. subjective vs. emotional content; by information source; by credibility; by time; by location; by embedded links; or by environmental relevance (physical, built or social). Automatic categorisation efforts have focused on supervised learning technologies, often backed by human annotation of crisis data [12]. Building on this work, Event Tracker provides fully automatic real-time reports alerting that are predicted to contain critical and actionable information, as well as state-of-the-art content-based categorisation based on a crisis event ontology (from TREC-IS, which we discuss below).

Crisis Content Processing and TREC-IS: As discussed above, a range of automatic technologies to support crisis management have previously been proposed. However, individually they have had little impact on crisis management as a whole [16]. One reason for this lack of impact is data quality [10]. To avoid a similar fate, Event Tracker's automatic content categorisation service targets a crisis ontology being maintained by an active initiative, namely TREC Incident Streams (http://trecis.org). TREC-IS is a new track at the Text Retrieval Conference (TREC) started in 2018 and designed to unite academia and industry around research into automatically processing social media streams during emergency situations and categorising information and aid requests made on social media for emergency service operators. TREC-IS provides training datasets for training crisis content categorisation systems, as well as an on-going (bi-yearly) evaluation challenge aimed at increasing the TRL of such systems.

Related Systems: Before discussing the design and architecture of Event Tracker, it is worth highlighting some related systems that we either build upon or learn from. First, *EAIMS* [13], was a prototype crisis management system that aimed at exploiting social media data to provide real-time detection of emergency events (demoed at SIGIR 2016), along with search, sentiment analysis, discussion-thread extraction and summarisation functionalities. The main drawback of EAIMS is that it was primarily designed only for use by high-level, head-office, emergency response staff, in contrast to Event Tracker that also targets volunteer groups. *Twitter Moments* is a relatively new feature of the Twitter platform, added in 2015. A 'Moment' is a user curated collection of Tweets, which allows users to comment on stories, promote news or create round-ups of different world events [2]. Twitter Moments has great potential for providing a high-level overview of an event. However, a key downfall is that Moments are popularity-driven, so local content such as individual calls for help will be missed. Instead, Event Tracker relies on recall-focused crawling in conjunction to automatic content categorisation to find actionable information quickly and accurately. Furthermore, *AIDR* is a platform developed at the Qatar Computing Research Institute that collects and classifies crisis-related Tweets [12]. AIDR makes use of crowdsourcing to label Tweets, which can then be used to train an automatic classifier. This ensures that the system performs well for each new event registered, since automatic classification using only pre-trained models can lead to low accuracy. Currently, Event Tracker relies on pre-trained classification models, however its underlying microservice architecture is flexible and allows for the integration of additional services like AIDR. Finally, *Twitcident* is another web-based system for filtering and analysing social media data on real-world incidents, such as disasters. Similar to EAIMS, Twitcident [1] provides incident detection and search technologies, enabling emergency services to filter data to aid their response efforts. However, a relatively recent study showed that Twitcident did not provide much value to response agencies in practice, as it could not provide early warnings regarding critical information [3]. For this reason, Event Tracker integrates criticality estimation tools to provide real-time reporting of crucial information.

3 EVENT TRACKER ARCHITECTURE

As argued earlier, the goal of Event Tracker is to provide an integrated platform that provides both automated low-latency ingestion and augmentation of report streams (either manually entered or crawled from social media) with effective support tools that enable response officers and volunteers to collaborate together to generate a unified operational picture during an emergency. In effect, this means that the platform must support: 1) low-level support for different report streams (news reports, social media and manual form-filling by call-centre operators), 2) integration of fast text processing technologies for real-time report tagging, as well as both 3) classical on-request information access (search and content look-up) and 4) continuous push notification servicing.

As such, Event Tracker uses a flexible architecture as illustrated in Figure 2. The lowest ingestion layer provides multi-stream integration, using a common data back-end to [13]. Above this sits the augmentation layer, which enables modular and scalable integration of report tagging microservices (in this case for identification of actionable information and information criticality analysis) using a distributed Apache Flink back-end. The output of the augmentation layer is then processed by the activity layer, that handles the 'business logic' for the application (data and user management), feeding either persistent storage structures (a search index and database) and/or the front-end directly via push notifications. Indeed, during a disaster event, the volume and rate at which relevant content that needs to be processed may vary greatly (anywhere from 10 posts/min to 4,000 posts/min [7]). Thus, the architecture behind the Event Tracker is designed to handle high-volume streams of data.

4 KEY FUNCTIONALITIES

Event Tracker integrates a number of functionalities following the design vision discussed earlier. Below, we summarise these main functionalities.

Generating Actionable Information Feeds: As discussed above, getting key information to the right people is crucial for a successful crisis management system, and one of the main reasons why previous systems have failed. Event Tracker combats this by providing multiple actionable information feeds through user dashboards, each of which is designed with a different intention of the value it can provide to the response agencies. Figure 1 (a) pictures the *Twitter Feed*, a real-time, filterable collection of all data gathered across the duration of a given crisis event. The *Media Feed*, displayed in

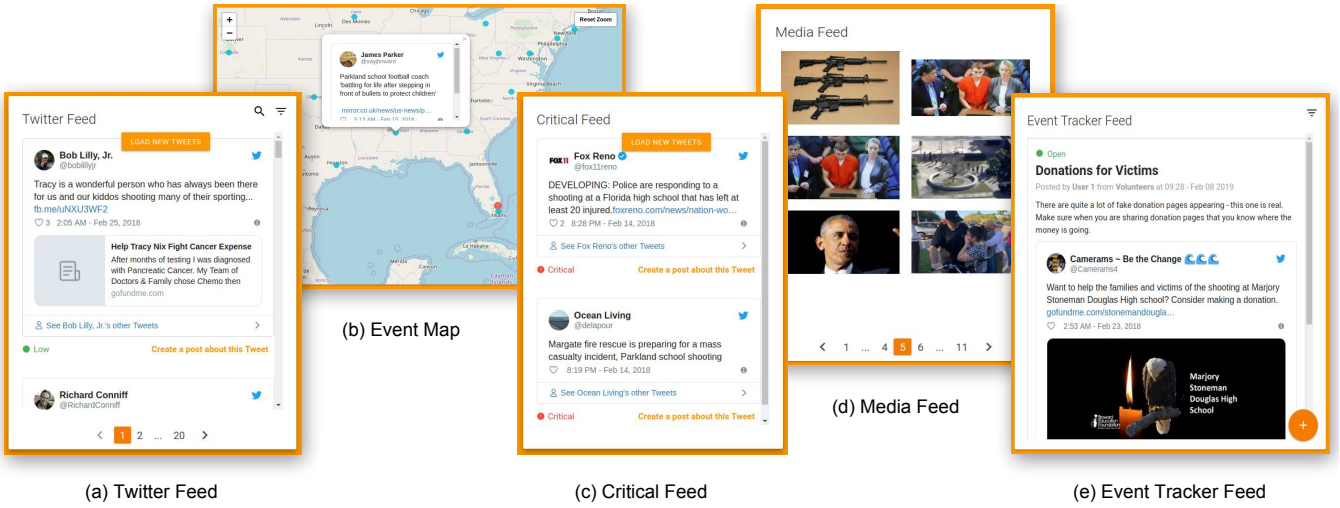

Figure 1: A selection of different functionalities available on the dashboards on Event Tracker

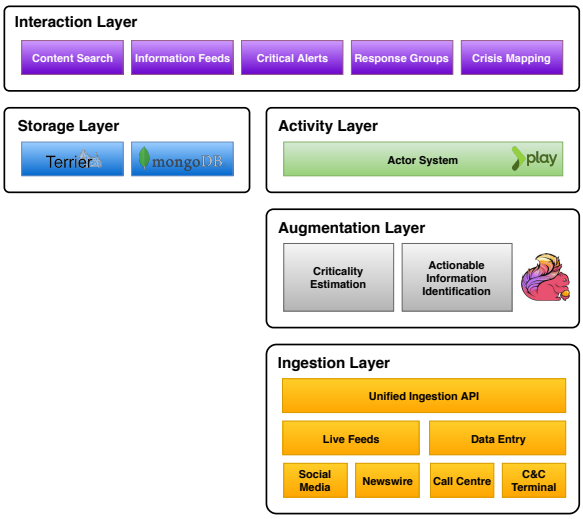

Figure 2: Conceptual architecture

Figure 1 (d), extracts all media collected, enabling response officers to, for example, assess damage during and after a disaster. Automatic content-based categorisation, through the TREC-IS initiative, examined in Section 2, is exploited by Event Tracker to provide emergency response agencies with a set of *Tweet Category Feeds*, in an attempt to increase how quickly and accurately these agencies can navigate the collection of data to find actionable information. Different agencies with different motivations can customise the set of feeds to distinctive categories such as 'First Party Observation' or 'Official Reports'. Each feed on Event Tracker aims to improve the situational awareness of the agencies, which is vital in effective decision making in complex and dynamic environments [8].

Criticality Alerting: Early warnings and alerts regarding critical information posted online during a crisis situation is imperative for a quick response from emergency services. Aid requests could be answered quickly, increasing for example the likelihood that a victim could be assisted in time. Event Tracker aims to accomplish this by also making use of the TREC-IS initiative to provide real-time identification of critical information. Fast classification and text processing technologies are exploited to immediately label incoming messages with a criticality score of *low*, *medium*, *high*, or *critical*, which can then be displayed alongside Tweets on a user's dashboard. Continuous push notifications servicing is also harnessed by Event Tracker to populate the *Critical Feed*, pictured in Figure 1 (c), to highlight reports which are predicted to contain critical and immediately actionable information.

Communicating Amongst Response Associations & Groups: To support crisis management, there are a large number of associations that could help monitor and manage a crisis (from the Red Cross to smaller Regional Volunteer Groups). Each event being tracked on the system can be related to multiple response associations, each of which may provide support in different manners. These associations can communicate through Event Tracker, using the *Event Tracker Feed* (Figure 1 (e)). Members of the response teams can use this feed to communicate directly with other response agencies, passing on any relevant information. As discussed, there has been a lack of successful systems that integrate formal response agencies with volunteers, and so for each event registered on Event Tracker, an open volunteering group is created, which anyone can join. The system can be used to increase the coordination and communication between these entities, with both volunteers on the ground reporting information, or digital volunteers highlighting critical reports to official response officers by embedding Tweets into their posts.

Additional Functionalities: Along with these key functionalities, Event Tracker provides some other features which are worth mentioning. First, we describe the *Crisis mapping* functionality.Figure 1 (b) pictures the *Event Map* module on Event Tracker, which displays both the location of the related event and the geo-tagged Tweets that have been collected over the duration of the crisis. As the information feeds, this aids in increasing the situational awareness of the

response agencies. The module could be extended in the future to display the locations of other reports, to improve data management. Next, we describe the *On-request information access* functionality. Making use of the Terrier search engine [14], Event Tracker allows users to explore the collected data with ease. New feeds are created on the user's dashboard, enabling multiple searches to run concurrently. For example, emergency response agencies can make use of this feature to discover reports relating to specific queries, such as 'fire'.

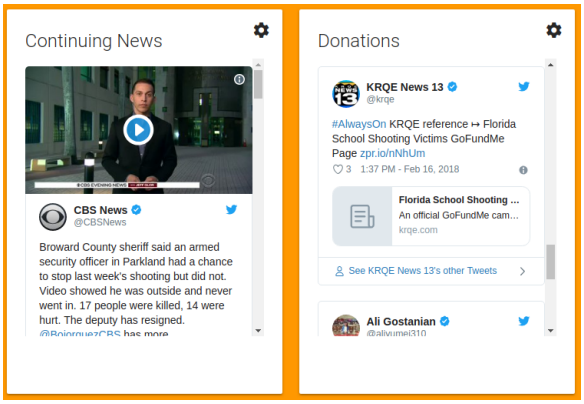

Figure 3: Potential Tweet Category feeds on user dashboards

5 ILLUSTRATIVE USE CASES

Focussing on multiple types of users during a crisis event, Event Tracker has several use cases. To illustrate the working of Event Tracker, let's consider two possible user cases. In the first use case, consider an emergency response officer, who uses the system to view all the actionable information feeds and navigate the incoming reports of data. During a simulation of the 2013 Australia Bushfire crisis, Event Tracker marked the Tweet "*Horrid emergency unfolding at #faulconbridge - fire is 50m behind #springwood shs - near hundreds of homes, norman lindsay gallery #nswfires*" as a high criticality emerging threat. Had Event Tracker been deployed over this event, a response officer accessing this report could have contributed an effective response. During an event, these officers may be monitoring multiple communication channels [17], which may lead to sporadic focus on each of the information feeds. As a consequence, some important information about tnhe event might be missed. Instead, with Event Tracker, the criticality alerting component would notify the user with any new reports that are predicted to be of vital importance, ensuring that the end-user can act quickly and efficiently. As a second case, consider a volunteer on the ground during a disaster. They could make use of Event Tracker to coordinate with response agencies to ensure that they are providing the best help they can. They can also make reports for the response agencies, to bring any significant information they possess to the attention of the response officers.

6 CONCLUSIONS

In this paper, we have presented Event Tracker, a modular and extensible prototype system designed to support the monitoring and management of events during crisis situations. The system leverages a flexible architecture, which is designed for low latency, and high-volume streams of data to provide functionalities such as actionable information feeds and criticality alerting. Event Tracker also enables the communication between various emergency response agencies and volunteers, and provides additional functionalities such as crisis mapping and on-request information access. In the short term, we aim to make Event Tracker available to the participants of the TREC-IS track. In particular, due to the modular augmentation layer Event Tracker is built upon, microservices such as those proposed by different groups participating to the TREC-IS track could be integrated and deployed, enabling different demonstrations to the end-users, whereby the corresponding outcomes and effectiveness of different technologies can be evaluated and further investigated.

REFERENCES

[1] Fabian Abel, Claudia Hauff, Geert-Jan Houben, Richard Stronkman, and Ke Tao. 2012. Twitcident: Fighting fire with information from Social Web streams. *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web Companion* (04 2012). https://doi.org/10.1145/2187980.2188035
[2] Liz Alton. 2018. Everything you need to know about Twitter Moments. https://business.twitter.com/en/blog/Everything-you-need-to-know-about-Twitter-Moments.html. Accessed: 09.02.2019.
[3] Kees Boersma, Dominique Diks, Julie Ferguson, and Jeroen Wolbers. 2016. *From Reactive to Proactive Use of Social Media in Emergency Response: A Critical Discussion of the Twitcident Project.* https://doi.org/10.4018/978-1-4666-9867-3.ch014
[4] Carlos Castillo. 2016. *Big Crisis Data: Social Media in Disasters and Time-Critical Situations.* Cambridge University Press. https://doi.org/10.1017/CBO9781316476840
[5] Carlos Castillo. 2016. *Big crisis data: social media in disasters and time-critical situations.* Cambridge University Press.
[6] Lise Ann St Denis, Amanda L Hughes, and Leysia Palen. 2012. Trial by fire: The deployment of trusted digital volunteers in the 2011 shadow lake fire. In *Proceedings of ISCRAM.*
[7] R. McCreadie et al. 2016. D4.7 - Integrated Search over Social Media. *Deliverable, SUPER FP7 Project* (2016).
[8] J. Harrald and T. Jefferson. 2007. Shared Situational Awareness in Emergency Management Mitigation and Response. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07).* 23–23. https://doi.org/10.1109/HICSS.2007.481
[9] Starr Hiltz, Jane Kushma, and Linda Plotnick. 2014. Use of Social Media by U.S. Public Sector Emergency Managers: Barriers and Wish Lists. https://doi.org/10.13140/2.1.3122.4005
[10] Starr Roxanne Hiltz, Jane A Kushma, and Linda Plotnick. 2014. Use of Social Media by US Public Sector Emergency Managers: Barriers and Wish Lists.. In *Proceedings of ISCRAM.*
[11] Peter Hoeppe. 2015. Trends in weather related disasters - Consequences for insurers and society. (2015).
[12] Muhammad Imran, Carlos Castillo, Ji Lucas, Patrick Meier, and Sarah Vieweg. [n. d.]. AIDR: Artificial intelligence for disaster response. In *Proceedings of WWW.*
[13] Richard McCreadie, Craig Macdonald, and Iadh Ounis. 2016. EAIMS: Emergency Analysis Identification and Management System. https://doi.org/10.1145/2911451.2911460
[14] Iadh Ounis, Gianni Amati, Vassilis Plachouras, Ben He, Craig MacDonald, and Christina Lioma. 2006. Terrier : A High Performance and Scalable Information Retrieval Platform. In *Proceedings of OSIR'2006.*
[15] Andrew Perrin. 2015. Social Media Usage: 2005-2015. Accessed: 16.02.2019.
[16] C Reuter, G Backfried, MA Kaufhold, and F Spahr. 2018. ISCRAM turns 15: A Trend Analysis of Social Media Papers 2004-2017. In *Proceedings of ISCRAM.*
[17] Andrea H Tapia, Kathleen A Moore, and Nichloas J Johnson. 2013. Beyond the trustworthy tweet: A deeper understanding of microblogged data use by disaster response and humanitarian relief organizations.. In *Proceedings of ISCRAM.*
[18] Marie Truelove, Maria Vasardani, and Stephan Winter. 2015. Towards credibility of micro-blogs: characterising witness accounts. *GeoJournal* (2015).
[19] J Twigg and Irina Mosel. 2017. Emergent groups and spontaneous volunteers in urban disaster response. *Environment and Urbanization* 29 (08 2017), 095624781772141. https://doi.org/10.1177/0956247817721413
[20] Joshua Whittaker, Blythe McLennan, and John Handmer. 2015. A review of informal volunteerism in emergencies and disasters: Definition, opportunities and challenges. *International Journal of Disaster Risk Reduction* 13 (2015), 358 – 368. https://doi.org/10.1016/j.ijdrr.2015.07.010