Didn't Roger That

Social Media Message Complexity and Situational Awareness of Emergency Responders

Pogrebnyakov, Nicolai; Maldonado, Edgar A.

Document Version
Accepted author manuscript

Published in:
International Journal of Information Management

DOI:
10.1016/j.ijinfomgt.2018.02.004

Publication date:
2018

Citation for published version (APA):
Pogrebnyakov, N., & Maldonado, E. A. (2018). Didn't Roger That: Social Media Message Complexity and Situational Awareness of Emergency Responders. *International Journal of Information Management*, *40*, 166-174. https://doi.org/10.1016/j.ijinfomgt.2018.02.004

Didn't Roger That: Social Media Message Complexity and Situational Awareness of Emergency Responders

Nicolai Pogrebnyakov and Edgar A. Maldonado

Journal article (Accepted manuscript*)

Please cite this article as:
Pogrebnyakov, N., & Maldonado, E. A. (2018). Didn't Roger That: Social Media Message Complexity and Situational Awareness of Emergency Responders. *International Journal of Information Management*, *40*, 166-174. DOI: 10.1016/j.ijinfomgt.2018.02.004

DOI: [10.1016/j.ijinfomgt.2018.02.004](https://doi.org/10.1016/j.ijinfomgt.2018.02.004)

Didn't roger that: Social media message complexity and situational awareness of emergency responders

Nicolai Pogrebnyakov[a,*] and Edgar Maldonado[b]

[a] Copenhagen Business School, 24 Porcelaenshaven, Frederiksberg 2000, Denmark
[b] Metropolitan State University of Denver, 1205 5th Street, Denver, CO 80204, United States
[*] corresponding author

This study investigates the role of social media in situational awareness in the emergency response domain. It builds a theoretical model to that effect, the first such effort to the best of our knowledge, and empirically investigates one of the components of the model, text complexity. The empirical analysis was performed on a dataset of 999,243 messages from 997 Facebook pages of US police departments in 2009—2016. Messages were classified into four categories based on their utilitarian or hedonic nature: emergency preparedness, emergency response, post-emergency and user engagement. Three measures of complexity were used, each capturing different aspects of text. Contrary to the hypothesis formulated in the study, messages in the post-emergency and the emergency response categories were found to be the most complex. With text complexity on social media being an underexplored area, these results suggest a need for an explicit study of the link between social media messages and situational awareness, and indicate a need for practitioners to revisit social media practices.

Keywords: social media; text complexity; situational awareness; emergency response

1 Introduction

The use of social media in emergency response has been gaining increased attention in recent years (Meier, 2015). Social media have been acknowledged to play a role at different stages of emergency response, from disaster response (Avvenuti, Cresci, Marchetti, Meletti, & Tesconi, 2016) to emergency preparedness (Merchant, Elmer, & Lurie, 2011), and in emergencies of different scale, from large-scale disasters such as earthquakes (Yates & Paquette, 2011) to smaller-scale emergency events, e.g. wildfires (Slavkovikj, Verstockt, Van Hoecke, & Van de Walle, 2014). In turn, the public increasingly expects emergency responders to communicate through social media (Lindsay, 2011).

In all but the simplest cases, emergency response involves several groups of actors (e.g. firefighters and police working side by side at the location where the emergency took place). Therefore, emergency-related information disseminated on social media by one actor (e.g., a police unit) is consumed by a diverse variety of other actors (e.g. other police units, firefighter and medical units and the public). In addition, emergency's responders understanding of the specific emergency influence how they produce and consume social media information. When looking for a theoretical construct that allows to frame social media messages within the context of emergency responders' awareness of a situation, the notion of situational awareness (SA) seems to be an appropriate choice. Situational awareness is a concept that the human factors community has been researching since the early 1990s (Endsley, 1995) and is the notion of "knowing what is going on so you can figure out what to do" (Yang, Chen, & Su, 2016).

All emergency responders need to attain situational awareness (SA) when dealing with a specific event, a phenomenon that has been labeled shared or intergroup SA (Seppänen, Mäkelä, Luokkala, &

Virrantaus, 2013; Sonnenwald & Pierce, 2000). SA, according to Endsley (1995), includes three stages: perception, comprehension, and projection. The first step, perception, deals with capturing the data. The comprehension stage has to do with the interpretation of that data, and the last step aims to predict situation's possible outcomes. During the perception stage, the subject captures data by means of the senses (seeing, listening, smelling, etc.). Some of the captured data can be written, and for that kind of data the comprehension degree will depend on text features such as length, content, structure, and readability (Jagtman, 2014).

One type of textual data contributing to SA is social media, which allow information about an ongoing emergency to diffuse quickly and provide a better view of post-disaster recovery efforts (Verma, et al., 2011; Yin, Lampert, Cameron, Robinson, & Power, 2012). To show how social media messages and actors' SA interact, and which of these interactions have been studied previously, we propose a simple model shown in Figure 1. It follows the categorization of social media messages that are sent and received during disasters developed by Reuter, et al. (2012). The categorization considers two types of actors: organizations involved in emergency response and general public, and the four possible combinations of messages between them (emergency responders to emergency responders, emergency responders to the public etc.). Social media influences SA of both groups, and their SA in turn impacts the crafting of new messages.

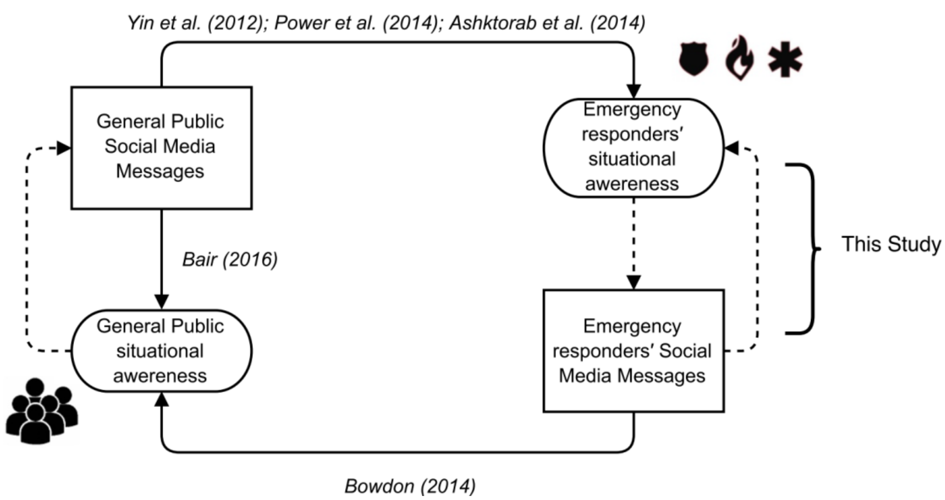


Figure 1. Social media messages and actors' situational awareness relationships.

The model includes previous work that can be categorized using information's flow. Dotted lines represent paths that have not been studied. There are three unexplored paths: how SA of the general public SA affects social media messages posted by it; the impact of messages posted by emergency responders on their SA; and the effect of SA of emergency responders on their messages.

This study focuses on the second path, how emergency responders' social media messages influences emergency responders' SA. Our approach revolves around the notion of complexity. Since emergency responders need to provide accessible information to other actors, communication should be done at a level that is readily understood by them, as these actors may not "speak the same language": e.g. some terms may be specific to a profession or location, and some acronyms may not be widely known. For this reason emergency-related communication should be simple, and calls to that effect have been issued by academics (Temnikova, Vieweg, & Castillo, 2015) and practitioners alike (International Association of Fire Chiefs, 2009).

However, whether emergency information on social media is indeed communicated in simple language in practice is not well understood. So far only very limited research has examined language simplicity in emergency-related social media communication, see Temnikova, et al.(2015). Our paper aims to contribute to this nascent field by comparing the degree of complexity of different types of information communicated by emergency responders on social media. And, rather than viewing text complexity as an end in itself, we use Endsley's (1995) theory of SA to develop a model where complexity affects SA among involved actors, which in turn may lead to actions to respond or adapt to the emergency.

Specifically, our goal is to empirically rank types of messages created by emergency responders by simplicity of language used in these messages. To achieve this goal, we collected and classified 999,243 Facebook messages of 997 US local police departments from January 2009 to October 2016. These messages were classified into four categories: emergency preparedness, emergency response, post-emergency and engagement with users. We used three measures of text complexity, which use different operationalizations of complexity. Results indicate that post-emergency and emergency response messages were the most complex using all three complexity measures, and that trend has been consistent throughout most of the period under consideration.

This research contributes to the underexplored area of social media message complexity in the emergency response realm, and the effect of emergency-related messages' complexity on SA. We develop a model of social media-assisted SA in emergency response and explore the relationship between social media message complexity and message type. With that information, emergency manager organizations can develop guidelines to facilitate the generation of effective social media communications that lead to better SA. Additionally, our model can be used by other researchers studying how different type of data (pictures, diagrams, etc.) influence emergency responders' SA.

2 Literature review

2.1 Situational awareness during emergencies

Situational awareness is defined as "all knowledge that is accessible and can be integrated into a coherent picture, when required, to assess and cope with a situation" (Sarter & Woods, 1991). SA has been extensively studied at the individual level, and in dynamic environments such as air traffic control (Jensen, 1997), aviation (Sarter & Woods, 1991), military (Sonnenwald & Pierce, 2000) and emergency response (Seppänen, et al., 2013). Research has highlighted factors that facilitate SA (e.g., information processing capabilities of an individual and their workload), as well as the effect of increased awareness on decision-making and taking action in response to the situation (Endsley, 1995).

Endsley (1995) developed a theory of SA where awareness is a mediating factor between, on the one hand, technical system factors such as system capabilities and complexity as well as individual factors, and on the other hand decision-making and taking action. Endsley (1995) points out how firefighting, police and military personnel depend on their SA to make decisions. Since social media messages are crafted depending upon specific circumstances, they are byproducts of a decision making process.

More recently, the concept has been found applicable to the group level as well. In fast-changing situations SA needs to be formed not only by individuals, but also within and between groups to facilitate efficient response to the situation (Nofi, 2000; Sonnenwald & Pierce, 2000). Specifically in the context of emergency response, different groups are likely to perform different roles in the response process: e.g., during a forest fire police, firefighters and medical services are performing different tasks. Each group thus needs to develop SA both within their own group to successfully perform those tasks, as well as a shared SA to not obstruct the performance of other groups.

Achieving SA, however, is not an easy task. It can be hampered, among others, by a lack of fluency in communication or by divergent understanding of some concepts by different actors (Seppänen, et al., 2013). Sonnenwald & Pierce (2000) report that in the military settings, actors observed that "we argue constantly over definition of terms". Therefore clarity in communication can be expected to facilitate SA.

Information technology has long been seen as one tool that could aid in increasing SA (Sonnenwald & Pierce, 2000). More recently, social media in particular have been shown to aid in increasing SA. Vieweg, et al. (2010) outlined features that can be extracted from tweets to increase SA in an emergency, and Verma, et al. (2011) developed a method to automatically identify tweets that can contribute to increasing SA during an emergency. Yin, et al. (2012) created a system that automatically processes tweets to detect sudden increase in activity, classify tweets by topic and visualize the data, with the goal of aiding emergency officers better understand the situation. However, research on the role of social media in SA in still in its infancy.

2.2 Social media

Social media are defined here as web-based services that allow users to create content, establish connections with other users and share content with other users (Treem & Leonardi, 2013). They offer several functionalities that are particularly relevant in the emergency response realm: visibility of information, establishment of connections and sharing of information and knowledge (Treem & Leonardi, 2013). Specific types of information that can be shared include text, images, videos and web links (Yates & Paquette, 2011).

Content created and shared on social media can be authored by individuals or organizational actors, both commercial and non-commercial. Early social media sites (e.g. MySpace) relied primarily on content generated by individuals, with researchers referring to these sites as "friend networking sites" (Bonds-Raacke, 2010; Fullwood, Sheehan, & Nicholls, 2009). However, recently the balance has been shifting in favor of content created by institutional actors. Facebook users in particular are posting less content, instead preferring to repost content of others, including professional content such as news articles (Griffith, 2016).

Social media content can be divided into two types: utilitarian and hedonic (Lin & Lu, 2011). Utilitarian content provides instrumental or productivity-enhancing value to users, such as maintaining relationships or obtaining new information (Piskorski, 2011; van der Heijden, 2004). Hedonic content, by contrast, is aimed at increasing users' level of enjoyment and pleasure (Lin & Lu, 2011; van der Heijden, 2004).

The idea of using social media to increase SA before, during, and after a crisis has been studied before (Watson & Rodrigues, 2017). Researchers and practitioners express a mostly positive view of social media impact on communications in the emergency field (Houston, et al., 2015). Nevertheless, emergency management agencies still find obstacles to adopting and using social media. Plotnick & Hiltz (2016) surveyed over 200 county level emergency managers and found that only about half of agencies use social media. Among the obstacles for social media adoption, the authors found that the lack of staff and guiding and policy documents were the main barriers for the use of social media. Plotnick & Hiltz (2016) further found that organizational and technical changes are needed to embrace social media, especially in the area of emergency management.

Additionally, social media also face the same challenges as other forms of communication during crisis. Fischer, et al. (2016) categorized the communication barriers during a crisis found in the literature. According to the authors, those barriers are technological, organizational, or social. In the social barriers category, one of the aspects has to do with the fact that "communication does not meet the requirements of the situation due to inadequate message design".

Social barriers especially resonate with social media. Social media is a conversational form of speech, "with multiple sources (...), varying levels of quality and grammatical correctness, and different languages present in the same corpus and sometimes in the same message" (Castillo, 2016). Effective design of social media messages is a field of study still in the early ages. Nevertheless, developing clear and concise messages is a critical component for effective communications during an emergency (Hyer & Covello, 2005).

The literature uses the terms " Crisis Informatics" and "Disasters Informatics" to describe the line of research that studies the use of information and communication technologies during emergencies (Anderson & Schram, 2011). This area of study justifies itself since, according to the World Health Organization, "effective media communication is in fact a crucial element in effective emergency management and should assume a central role from the start" (Hyer & Covello, 2005). Even more, communication is a critical component during all the phases related to an emergency (Houston, et al., 2015). Also, the use of new technologies during emergencies such as low-power wireless technologies, sensors, and ubiquitous connectivity give more reasons to pay attention to the novel field of Crisis Informatics.

Given the arrival of mobile devices, which are often used for interacting with social media (Castillo, 2016), it seems imperative to pay attention to how social media communication flows during a crisis. They are already institutions monitoring social media communications to improve their operations. For example, in 2014, the American Red Cross created its Media Digital Operations Center for Humanitarian Relief, which focuses, among other things, on sourcing information from areas affected by emergencies as well as connecting individuals with resources they need (e.g., food, shelter or emotional support) (American Red Cross, 2014).

Several phases of emergencies are typically identified: 1) mitigation (activities aimed at increasing resilience and decreasing vulnerability to disasters, e.g. zoning, barrier construction); 2) preparedness (creating specific emergency response plans, public education, staff training); 3) response (actions in a short period before, during and after the disaster aimed at addressing immediate needs of the affected population); and 4) recovery (longer-term activities to repair property and restore communities) (Altay & Green, 2006; Perry, Lindell, & Tierney, 2001).

The first step in analyzing social media messages is to collect them. One of the main issues with this task is how convoluted and vast is the pool of messages that need to be analyzed. Social media venues generate immense amounts of data during an emergency. So much data, that its processing exceeds human capacity (Castillo, 2016). This mass of data presents a challenge for practitioners and researchers: there is valuable and critical information, but there is also unverified and incomplete information from unknown sources (Howe, Jennex, Bressler, & Frost, 2011; St Denis, Palen, & Anderson, 2014).

2.3 Readability and complexity on social media

Now, given the diverse form of content that can be distributed by social media, the way emergency organizations use them will not follow the same methods that those use with traditional media sources. Nevertheless, one of the pillar principles of communication is still applicable: "communications need to be simple" (Hyer & Covello, 2005). Simple messages are needed for both interorganizational flow of information and for information from organizations to the public, which may perceive messages better if they are simple.

Seppänen, et al. (2013) underscore the need for fluent communication to adequately form SA. Communication fluency in turn is determined, among other factors, by the use of common concepts understood by all actors. Sonnenwald & Pierce (2000) suggest that the ability of actors to express their understanding of the situation and their intent clearly as one of the critical components needed to establish and maintain shared SA. Jones & Endsley (1996) identify several factors that may result in an actor to not perceive the situation correctly, thereby negatively affecting SA. Among them is the misperception of data that the actor attended to, which accounted for 8.7% of errors related to SA in 143 aviation incidents studied in that research. To address this, Stanton, et al. (2001) propose to present information in such a way that would make understanding the situation easier, thereby positively affecting SA. Finally, Nofi (2000) lists, again, misperception of information, along with "perception conflict" where some actors perceive information differently than others (due to e.g. differences in interpretation), as factors that degrade group SA. In short, much of extant research points to the need for greater simplicity in communication and clearer language as facilitators of SA.

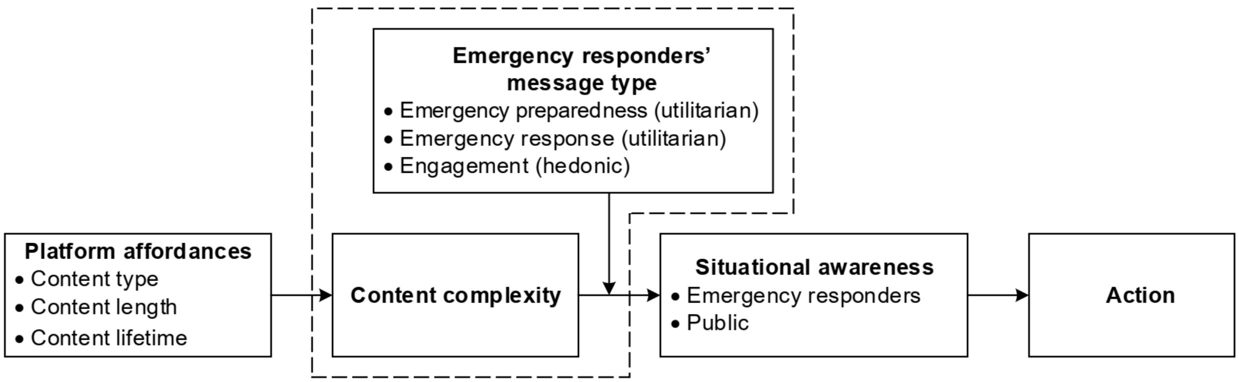
Practitioners have also called for an increased use of simpler language in communications about emergencies (International Association of Fire Chiefs, 2009). The purpose is to enhance cooperation and promote SA between different types of emergency responders (e.g. firefighters and police), as well as aid the public in understanding the emergency situation. While that particular case referred to radio communications, it stands to reason that emergency communication on social media should also strive for "plain language", with greater adoption of social media by emergency responders and the increased expectation from the public that information about emergencies is communicated through this channel (Ma & Yates, 2014).

Several studies have examined complexity of social media communications (albeit very rarely in the emergency response domain). Mitkov & Stajner (2014) developed a set of rules (e.g. "use simple sentences", "only use active voice") that are aimed at simplifying text. Risius & Pape (2015) acknowledge the popularity of the Flesch score, which is one measure of text complexity, and propose using another measure, the New Dale-Chall Readability formula to gauge the complexity of Twitter messages. Temnikova, et al. (2015) review and identify text characteristics that affect readability of Twitter messages, such as message length, the use of abbreviations, misspellings etc.

Despite these efforts, there is not a significant amount of literature on the study on how simple or readable are social media messages during an emergency (Temnikova, et al., 2015). However, the idea that clear and concise messages are critical for effective media communication is accepted by the emergency relief community (Hyer & Covello, 2005).

2.4 Theoretical model

Weaving the above discussed strands of literature together and building on Endsley's (1995) theory of factors conditioning situational awareness, we contextualize SA in the emergency response domain and apply it to social media technologies. Because we are interested in shared situational awareness, we focused on the system and the situational awareness level in Endsley's theory, and exclude the individual level. This resulted in the framework shown in Figure 2.



Note: dotted area represents the empirical focus of this study.

Figure 2. Model of social media-aided situational awareness in emergency response (adapted from Endsley, 1995).

Endsley's factors related to system capabilities were translated to the construct of "platform affordances", which describe any limitations the social media platform places on the type of content that can be posted there (e.g. video in the case of YouTube), content length (Twitter's character limit) or the lifetime of the content (self-destructing messages on Snapchat). While Endsley (1995) describes system limitations due to technological constraints, the affordances construct better reflects the conscious choices platform designers put into limiting these content aspects, rather than inherent technological constraints that prevent from posting content of e.g. specific types or length.

Platform affordances in part dictate the level of complexity of the content that could be posted on the platform. Content complexity in turn affects SA (Seppänen, et al., 2013; Sonnenwald & Pierce, 2000).

The link between complexity and SA is moderated by the type of the message, which could be utilitarian or hedonic (Lin & Lu, 2011). Contextualizing utilitarian content, we differentiate between phases of emergency. We focus on the emergency preparedness, response and post-emergency recovery categories (and disregard the mitigation phase, as it involves long-term activities such as zoning, insurance that have a low impact on SA, and in practice distinguishing social media messages between mitigation and preparedness stages would prove difficult). It is likely that the conditions surrounding each phase (e.g. the pace of change, the number of involved actors) affects how content is perceived and processed, and in turn impacts SA. We also include hedonic content: on social media these are messages that emergency responders post in "quiet times" between emergencies to engage with the public, and we label these messages as "engagement".

Hedonic content seems to be a driver for general public participation in social media (Verma, Jahn, & Kunz, 2012). Nevertheless, giving the "entertainment" character of hedonic messages, we can argue they will not contribute on the decision process that characterizes the SA process. This type of content might contribute to get more listeners from the general public domain, but we can argue that it unlikely will impact emergency responders' SA.

The final step in Endsley's theory, which is also reflected in our model, is action taken based on SA.

In this study, we are specifically interested in the mediating role played by the emergency phase on complexity. In other words, does complexity of content published by emergency responders on social media differ between emergency phases and if so, how? The discussion above leads us to hypothesize that messages in the fast-paced emergency response phase should have a simpler language to facilitate SA. However, this needs to be empirically tested.

3 Methodology

3.1 Data

We collected and analyzed messages from Facebook profiles of US police departments. First, a list of URLs of police departments was collected from the websites usacops.com and policeapp.com. Each URL was then accessed and, if it had a link to the department's Facebook profile, the link to the profile was retained. In this way, Facebook profiles of 1,261 police departments were obtained. For each profile, messages in the period from January 1, 2009 to November 1, 2016 were obtained. After removing Facebook pages that have not posted a single message, the dataset used in the analysis included 997 pages and 999,243 messages.

Data analysis consisted of two steps. In the first step, a small sample of messages was used to train a classifier that could label a given message as belonging to one of the four categories: emergency preparedness, response, post-emergency and engagement. In the second step, that classifier was used to automatically categorize the remaining messages, and their readability scores were analyzed. The following two sections explain each step in detail.

3.2 Training the message classifier

Manual classification. To provide training data for the classifier, a random selection of 5,000 messages was manually classified into one of four categories shown in Table 1.

Recurrent neural networks (RNN). RNN is a state-of-the-art type of artificial neural networks that has been successfully used for automatic translation, sentiment analysis and text categorization, among others (Kim, 2014; LeCun, Bengio, & Hinton, 2015; Yogatama, Dyer, Ling, & Blunsom, 2017). It performs well on inputs that can be represented as a sequence, such as time series data or words in a text. RNN processes these inputs one by one, considering at each step not only the new input but also an internal parameter (called state) computed in the previous step using that step's input. This allows RNN to learn dependencies between inputs, such between words in a text when classifying that text.

RNN demonstrated the best performance among several classifiers on this dataset, including convolutional neural networks and support vector machines (SVM). Further details on classifier selection for this dataset are described in Pogrebnyakov and Maldonado (2017).

Training. The manually classified messages were transformed into feature vectors, split into a training and test datasets (see Table 2 for the number of examples in each set), and used to train a RNN classifier.

Message category	*Description*	*Examples*
Emergency preparedness	Preparedness information, tips on how to prepare for and behave in an emergency	"Watch out for ice and falling snow." "Halloween is just around the corner. Here are some safety tips for you and your children to go over before you purchase their costumes and before you go out trick or treating! [URL]" "Did you know you can text crime tips to CrimeStoppers? It's completely anonymous and tips leading to an arrest can pay up to $1,000!"
Emergency response	Update about an ongoing emergency	"Road Closure/2000 Q ST. NW /20th ST NW due to a gas leak in the area." "Robb/Gun- 3000 Blk 1st, SE. L/O Burgundy Honda ,B/M, blk shirt, bl/jeans, possibly blk shoes, dreads, S/2 blk shirt, blue jeans, ski mask." "If you have information about any of these individuals, please contact the Sheriff's Office at [phone]. Do not approach or attempt to apprehend."
Post-emergency	Tips for mitigating effects of a past emergency, new information about a past emergency	"The demonstration at Saint Paul and Lexington Street has disbanded. No traffic was impacted." "a suspect has been identified in the two housebreaks that occurred on Tuesday in Danvers and was linked to a car break and housebreak in Peabody the same day."
Engagement	Updates about the department's internal non-emergency operations, conversation with Facebook users	"Yesterday, the department formally recognized the Records Division for their hard work all year long. Next week all of the San Mateo County police records clerks will celebrate together at their annual recognition event. Pictures to follow!" "Check out today's Chronicle!! Commander [name] is featured as the "Most Admired Woman in Government." Congratulations Commander!!"

Table 1. Categorization of Facebook messages of police departments.

RNN architecture used in this classifier consisted of a single Gated Recurrent Unit (GRU) cell unrolled over 100 time periods and a softmax output layer. GRU helps overcome performance problems and prevent RNN's state from decaying or growing exponentially (exploding/vanishing gradients) when learning longer-term dependencies (Bengio, Simard, & Frasconi, 1994; Graves, 2012). The learning rate was 0.001 and a minibatch size of 50 was used.

Features were created from the text in input messages using word2vec embeddings (Mikolov, Chen, Corrado, & Dean, 2013). Word2vec represents input words as dense vectors that capture semantic and morphological similarities between words (e.g., "road" and "street" have similar meanings and are both nouns). These vectors need themselves to be learned first from raw text. While there are pretrained vectors available (e.g., ones trained on a Google News dataset of 100 billion words), we found that vectors trained on our entire dataset of messages worked best for this classifier, and used these embeddings. We represented each message using up to 100 first words in the message, and 84.4% messages in the dataset were shorter than that.

	Emergency preparedness	*Emergency response*	*Post-emergency*	*Engagement*
Training set	573	364	931	1,627
Test set	266	139	389	711

Table 2. The number of examples in the training and test sets, by message classes.

Classification. Since the classes are skewed as shown in Table 2, accuracy is not a good measure of the classifier's performance. Instead we use the F1 measure averaged across all classes (see Table 3) (Y. Yang & Liu, 1999). This yielded F1 value of 0.839, which compares favorably to other studies on classifying social media messages (Yu & Kwok, 2011). The classifier was then used to categorize the entire dataset of Facebook messages.

	Emergency preparedness	*Emergency response*	*Post-emergency*	*Engagement*	*Average*
F1 measure	0.776	0.785	0.806	0.902	0.839

Table 3. F1 measures: by class and average.

3.3 Analyzing message readability

Having classified each message in the dataset with the classifier, we calculated readability scores for each message in the dataset. Three readability scores were used to assess the complexity of a message: the Flesch reading ease score, FRE (Flesch, 1948), the Automated readability index, ARI (Smith & Senter, 1967) and the New Dale-Chall Readability formula, DCR (Chall & Dale, 1995). We chose these measures because they assess complexity from different standpoints: as the length of words and sentences (in the case of ARI), including the number of syllables per word (FRE), and considering widely used words in the case of DCR.

Specifically, the Flesch reading ease score (Flesch, 1948) is calculated as:

$$\mathrm{FRE} = 206.835 - 846 \cdot \frac{\mathrm{SyllablesCount}}{\mathrm{WordsCount}} - 1.015 \cdot \frac{\mathrm{WordsCount}}{\mathrm{SentencesCount}}$$

Smaller values correspond to more complex text.

The formula for the Automated readability index (Smith & Senter, 1967) is:

$$\mathrm{ARI} = \frac{\mathrm{WordsCount}}{\mathrm{SentencesCount}} + 9 \cdot \frac{\mathrm{CharactersCount}}{\mathrm{WordsCount}}$$

Here, smaller values correspond to simpler text.

The New Dale-Chall Readability score (Chall & Dale, 1995) is calculated as:

$$\mathrm{DCR} = \begin{cases} 0.1579 \cdot \frac{\mathrm{HardWords}}{\mathrm{TotalWords}} - 0.0496 \cdot \frac{\mathrm{WordsCount}}{\mathrm{SentencesCount}}, \text{ if } \frac{\mathrm{HardWords}}{\mathrm{TotalWords}} \leqslant 0.05 \\ 3.6365 + 0.1579 \cdot \frac{\mathrm{HardWords}}{\mathrm{TotalWords}} - 0.0496 \cdot \frac{\mathrm{WordsCount}}{\mathrm{SentencesCount}}, \text{ if } \frac{\mathrm{HardWords}}{\mathrm{TotalWords}} > 0.05 \end{cases}$$

where HardWords is the number of words in the document which are not contained in the list of 3,000 "simple" words specified by the authors.

We used the Stanford CoreNLP package to obtain the number of syllables, words and sentences (Stanford NLP Group, 2017).

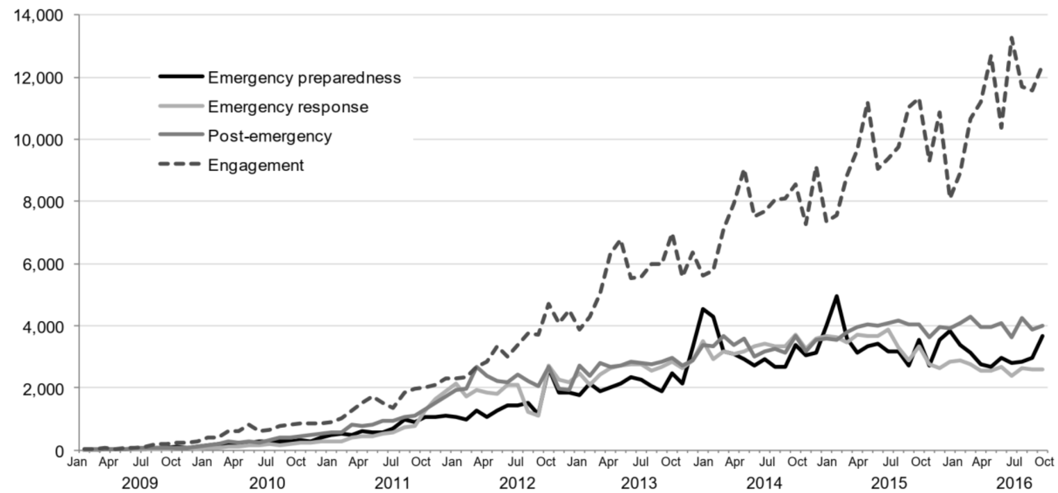
An interesting question could be whether these readability scores are applicable to social media messages, Facebook messages in particular, and this research. Some social media platforms (e.g., Twitter) limit the size of messages that could be posted on the platform. The formulation of readability scores does not consider the size of text being analyzed for readability. Thus message size in itself is not an obstacle for deploying these readability scores on social media messages. A corollary of the size limit, however, is that some users abbreviate words, or use hashtags where several words are collated (e.g. "#BeautifulDay"). Without special preprocessing these scores may report such messages as less readable, as the scores consider words with a greater number of syllables, or those not on the list of "easy" words, as more complex text. However, as Facebook does not limit the size of messages, this is unlikely to be a significant problem for this platform. Further, this research is interested in relative scores across message classes drawn from a single source, rather than comparing the readability of e.g. Facebook messages and journal articles. Therefore the choice of readability scores is appropriate for this research.

To ease understanding of these scores, we converted each of them to a scale from 0 to 1, with 0 corresponding to easy and 1 to complex text (thus the Flesch reading score was also inverted). The converted scores are denoted cFRE, cARI and cDCR respectively. Note that even though the converted scores are on the same scale, their values are not directly comparable across different scores as they have different mean values.

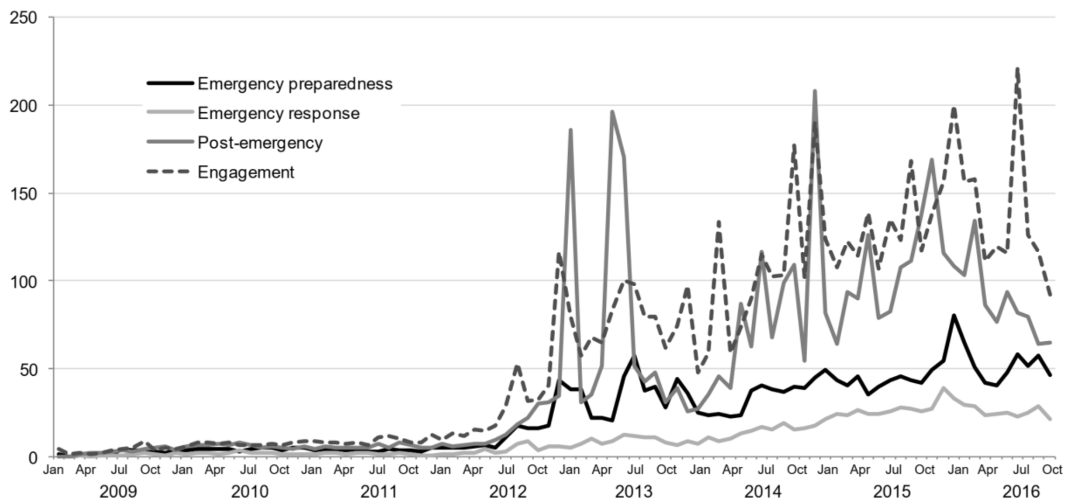
4 Results

4.1 Descriptive measures over time

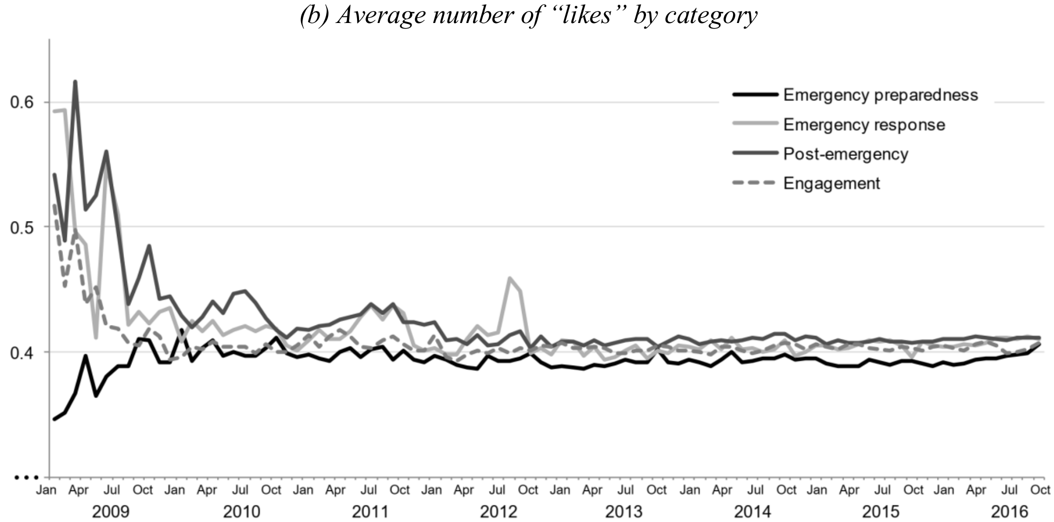
Figure 3 shows the number of messages, the average number of "likes" and average readability scores of messages in the dataset by month.



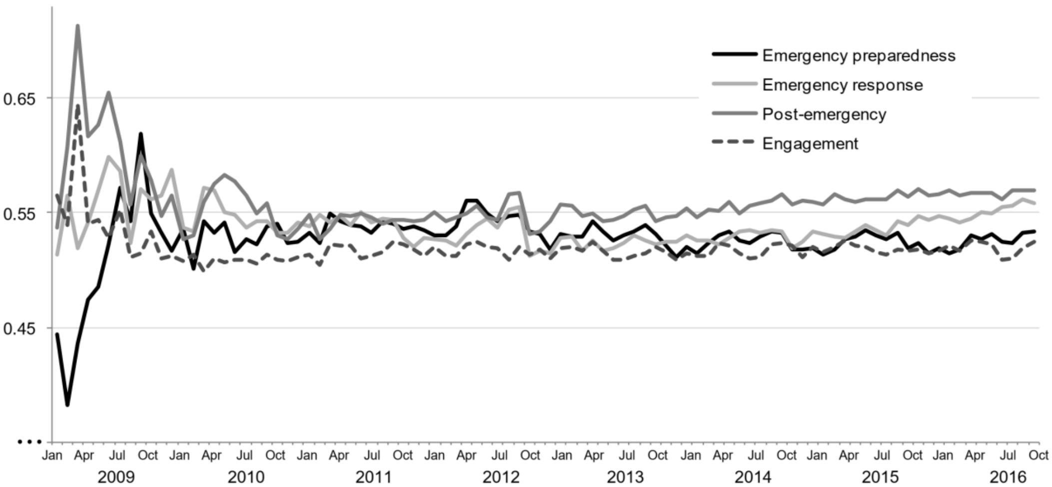
(a) Number of messages by category



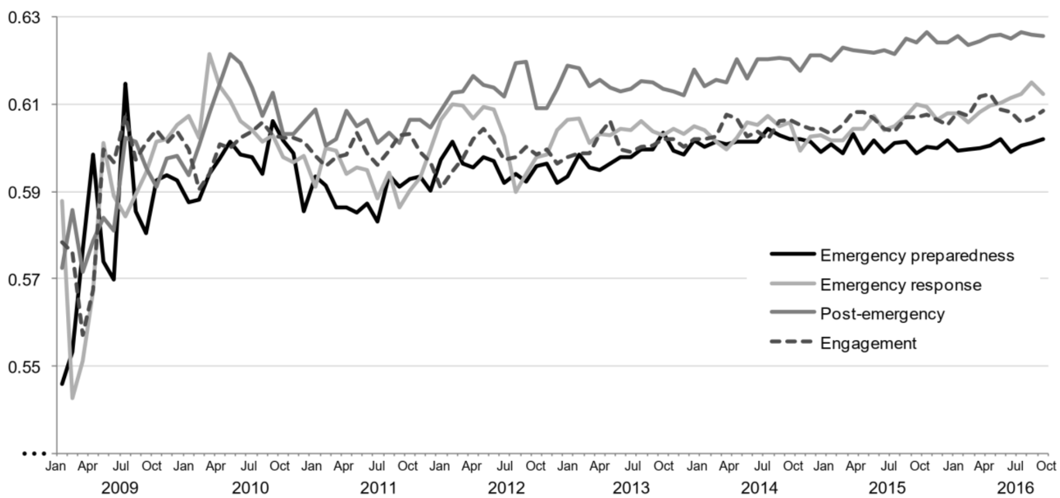
(b) Average number of "likes" by category



(c) Average cFRE score by category



(d) Average cARI score by category



(e) Average cDCR by category

Figure 3. Number of messages, average number of "likes" and average readability scores of Facebook messages of US police departments by month, January 2009—November 2016.

The number of messages has been steadily growing since 2009, with 22,639 messages posted on all accounts included in our dataset in October 2016 (Figure 3a). Messages in the engagement category were the most numerous, with post-emergency being second numerous, followed by emergency preparedness and response. At the same time, messages in the engagement category repeatedly received the most "likes" (Figure 3b). Since engagement messages are expected to generate more interaction, which includes "likes", this supports the validity of classification of these messages, even though the classification did not include any metric of user engagement.

Looking at the measures of complexity, emergency response messages are the most complex using all three scores: cFRE, cARI and cDCR (Figure 3c—e). According to the cFRE measure post-emergency messages have been the most complex throughout most of the period. The cARI measure also shows that emergency response messages have been the most complex most of the time, and on the cDCR measure response messages have maintained an uninterrupted highest level of complexity since late 2011.

4.2 Statistical analysis

Table 4 shows descriptive statistics of messages by category. Most messages in the dataset are in the engagement category, followed by post-emergency, emergency response and preparedness. Emergency response messages, on average, are the shortest, with the median message in that category having 22 words.

	Emergency preparedness	*Emergency response*	*Post-emergency*	*Engagement*
Number of messages	167,091	172,828	203,447	455,877
Median number of words per message	26	22	26	25
Mean cFRE score*	0.393 (0.082)	0.406 (0.104)	0.411 (0.082)	0.403 (0.092)
Mean cARI score*	0.527 (0.121)	0.533 (0.106)	0.557 (0.109)	0.517 (0.117)
Mean cDCR score*	0.599 (0.058)	0.605 (0.050)	0.618 (0.052)	0.604 (0.058)

*Standard deviation in parentheses.

Table 4. Descriptive statistics of different categories of Facebook messages in the dataset.

All four categories have high median FRE scores, which, at above 100, indicate messages that are easy to read.

These results do not support our hypothesis. Preparedness and engagement messages on average have higher readability than messages in the post-emergency and emergency response categories. This is demonstrated by greater mean values for these message categories using each of the three readability measures: cFRE, cARI and cDCR. Furthermore, the differences in mean values of readability scores between each of the messages categories are statistically significant, as shown in Table 5.

Readability measure	*ANOVA F value*
cFRE	1,489.8***
cARI	5,992.6***
cDCR	4,725.8***

***$p < 0.001$

Table 5. F values from the ANOVA tests for the equality of mean values between message categories.

4.3 Summary

Data analysis reveals that post-emergency and emergency response messages were the most complex using each measures of text complexity we used, and this pattern has persisted for most of the timeframe in our analysis. This difference was statistically significant. However, despite this relative difference, in absolute terms an average emergency response message was easy to read, as indicated by a high FRE score.

5 Discussion and conclusion

Social media have been recognized to have the potential to enhance SA (Vieweg, et al., 2010), and one of the factors affecting SA is message complexity. Simple and clear language is beneficial in communication (Mitkov & Štajner, 2014; Temnikova, et al., 2015). The importance of simplicity is arguably even greater in emergency communications, where multiple actors are often involved and response time is essential. Despite this importance, message complexity is a relatively underexplored area in social media research, and highly so in the emergency response domain. Extant results indicate that simpler language in messages has positive effect on organizational outcomes in e-commerce (Risius & Pape, 2015). However, to the best of our knowledge, ours is a first large-scale study of language complexity in the emergency response domain.

The theoretical contribution made in the article is a theoretical link between the construct of situational awareness and the social media artifact. Extant studies have done so empirically (Verma, et al., 2011), however, to the best of our knowledge this is the first effort to do this theoretically. Our framework expands on the idea of situational awareness and includes text complexity as a critical element on the SA process. This theoretical construction can serve as a tool to explore how other type of social media messages (images, videos, etc.) play a role in the decision making process of emergency responders.

Contrary to the hypothesized relationship, we found that in emergency response communication, social media messages that are related specifically to emergency response are not written in the most simple language. This is despite calls for greater clarity of language in emergency communications by both academics and practitioners (International Association of Fire Chiefs, 2009; Temnikova, et al., 2015).

The validity of our findings rests on the efficacy of the complexity scores used to develop our analysis. Since we have followed the generally accepted procedures to handle and manipulate the data, this study exhibits internal validity. The external validity for our analysis could be part of future research, which can replicate the procedures in this study using data from different countries and compare the results. Future research can also deploy qualitative methods (e.g., interviews or participant observation of emergency responders) to probe into reasons behind differing levels of complexity of social media messages in this domain, and possible ways of simplifying them.

While we did not explore the explicit link between message complexity and SA because we only empirically measured message complexity but not SA, extant studies have established complexity as a factor in SA (Endsley, 1995). Our findings highlight the need to explore this relationship in greater depth, probing into why emergency response and recovery messages tend to be the most complex as well as identifying the consequences of this for SA.

At the practical level, the results suggest that those involved in communication in dynamic, fast-changing environments with multiple actors (e.g., emergency response, military) should consider message complexity when disseminating content, including on social media. The goal would be that messages could be understood by a wide variety of actors and to reduce chances of misinterpretation. This in turn would increase SA and enhance cooperation between emergency responders. In "quiet times" between emergencies, actors should devote attention to developing communication guidelines, for example creating pre-constructed message templates for different types of situations (e.g. floodings, fires) similar to "amber alerts" in North America. Researchers can also use this study's analytic tools to create systems that measure messages' complexity on real time and provide immediate feedback to emergency responders writing them.

References

Altay, N., & Green, W. G. (2006). OR/MS research in disaster operations management. *European Journal of Operational Research, 175*, 475-493.

American Red Cross. (2014). The American Red Cross and Dell Launch First-Of-Its-Kind Social Media Digital Operations Center for Humanitarian Relief. http://www.redcross.org/news/press-release/The-American-Red-Cross-and-Dell-Launch-First-Of-Its-Kind-Social-Media-Digital-Operations-Center-for-Humanitarian-Relief Accessed February 2017.

Anderson, K. M., & Schram, A. (2011). Design and implementation of a data analytics infrastructure in support of crisis informatics research: NIER track. In *2011 33rd International Conference on Software Engineering (ICSE)* (pp. 844-847).

Ashktorab, K., & al., e. (2014). Tweedr: Mining twitter to inform disaster response. In *Proc. of ISCRAM*.

Avvenuti, M., Cresci, S., Marchetti, A., Meletti, C., & Tesconi, M. (2016). Predictability or Early Warning: Using Social Media in Modern Emergency Response. *IEEE Internet Computing, 20*, 4-6.

Bair, A. (2016). From Crisis to Crisis: A Big Data, Antenarrative Analysis of How Social Media Users Make Meaning During and After Crisis Events. In: Utah State University.

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks, 5*, 157-166.

Bonds-Raacke, J. (2010). MySpace and Facebook: Identifying Dimensions of Uses and Gratifications for Friend Networking Sites. *Individual Differences Research, 8*, 27-33.

Bowdon, M. A. (2014). Tweeting an Ethos: Emergency Messaging, Social Media, and Teaching Technical Communication. *Technical Communication Quarterly, 23*, 35-54.

Castillo, C. (2016). *Big Crisis Data*. Cambridge, UK: Cambridge University Press.

Chall, J. S., & Dale, E. (1995). *Readability revisited: the new Dale-Chall readability formula*: Brookline Books.

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors, 37*, 32-64.

Fischer, D., Posegga, O., & Fischbach, K. (2016). Communication Barriers in Crisis Management: A Literature Review. In *Twenty-Fourth European Conference on Information Systems (ECIS)*.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*, 221-233.

Fullwood, C., Sheehan, N., & Nicholls, W. (2009). Blog Function Revisited: A Content Analysis of MySpace Blogs. *CyberPsychology & Behavior, 12*, 685-689.

Graves, A. (2012). *Supervised Sequence Labelling with Recurrent Neural Networks* (Vol. 385). Berlin: Springer.

Griffith, E. (2016). Facebook Users Are Sharing Fewer Personal Updates and It's a Big Problem. http://fortune.com/2016/04/07/facebook-sharing-decline/ Accessed February 2017.

Houston, J. B., Hawthorne, J., Perreault, M. F., Park, E. H., Goldstein Hode, M., Halliwell, M. R., Turner McGowen, S. E., Davis, R., Vaid, S., McElderry, J. A., & Griffith, S. A. (2015). Social media and disasters: a functional framework for social media use in disaster planning, response, and research. *Disasters, 39*, 1-22.

Howe, A. W., Jennex, M. E., Bressler, G. H., & Frost, E. (2011). Exercise 24: Using Social Media for Crisis Response. *International Journal of Information Systems for Crisis Response and Management (IJISCRAM), 3*, 36-54.

Hyer, R. N., & Covello, V. T. (2005). Effective Media Communication during Public Health Emergencies. In. Geneva, Switzerland: World Health Organization.

International Association of Fire Chiefs. (2009). "Plain language" Use in Radio Communications. https://www.iafc.org/docs/default-source/uploaded-documents/iafcposition-plainlanguage.pdf Accessed February 2017.

Jagtman, H. M. (2014). Writing short alarm messages: A matter of education, training and practice. In R. D. J. M. Steenbergen, P. H. A. J. M. van Gelder, S. Miraglia & A. C. W. M. Vrouwenvelder (Eds.), *Safety, Reliability and Risk Analysis: Beyond the Horizon* (pp. 411-419). New York, NY: CRC Press.

Jensen, R. S. (1997). The Boundaries of Aviation Psychology, Human Factors, Aeronautical Decision Making, Situation Awareness, and Crew Resource Management. *The International Journal of Aviation Psychology, 7*, 259-267.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space and Environmental Medicine, 67*, 507-512.

Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436-444.

Lin, K.-Y., & Lu, H.-P. (2011). Why people use social networking sites: An empirical study integrating network externalities and motivation theory. *Computers in Human Behavior, 27*, 1152-1161.

Lindsay, B. R. (2011). Social media and disasters: current uses, future options, and policy considerations. *Congressional Research Service*.

Ma, X., & Yates, J. (2014). Optimizing social media message dissemination problem for emergency communication. *Computers & Industrial Engineering, 78*, 107-126.

Meier, P. (2015). *Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response*: CRC Press.

Merchant, R. M., Elmer, S., & Lurie, N. (2011). Integrating social media into emergency-preparedness efforts. *New England Journal of Medicine*, *365*, 289-291.

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

Mitkov, R., & Štajner, S. (2014). The Fewer, the Better? A Contrastive Study about Ways to Simplify. *COLING 2014*, 30.

Nofi, A. A. (2000). Defining and measuring shared situational awareness. In. Alexandria, VA: Center for Naval Analyses.

Perry, R. W., Lindell, M. K., & Tierney, K. J. (2001). *Facing the Unexpected: Disaster Preparedness and Response in the United States*. Washington, DC: Joseph Henry Press.

Piskorski, M. J. (2011). Social strategies that work. *Harvard Business Review, 89*, 116-122.

Plotnick, L., & Hiltz, S. R. (2016). Barriers to use of social media by emergency managers. *Journal of Homeland Security and Emergency Management, 13*, 247–277.

Pogrebnyakov, N., & Maldonado, E. (2017). Identifying emergency stages in Facebook posts of police departments with convolutional and recurrent neural networks and support vector machines. In *Proceedings of the BSMDMA-SocialNLP Workshop at IEEE Big Data*. Boston, MA, December 11-14.

Power, R., Robinson, B., Colton, J., & Cameron, M. (2014). Emergency Situation Awareness: Twitter Case Studies. In C. Hanachi, F. Bénaben & F. Charoy (Eds.), *Information Systems for Crisis*

Response and Management in Mediterranean Countries (pp. 218-231): Springer International Publishing.

Reuter, C., Marx, A., & Pipek, V. (2012). Crisis management 2.0: Towards a systematization of social software use in crisis situations. *International Journal of Information Systems for Crisis Response and Management (IJISCRAM), 4*, 1–16.

Risius, M., & Pape, T. (2015). Developing and Evaluating a Readability Measure for Microblogging Communication. In *E-Life: Web-Enabled Convergence of Commerce, Work, and Social Life* (pp. 217-221): Springer, Cham.

Sarter, N. B., & Woods, D. D. (1991). Situation Awareness: A Critical But Ill-Defined Phenomenon. *The International Journal of Aviation Psychology, 1*, 45-57.

Seppänen, H., Mäkelä, J., Luokkala, P., & Virrantaus, K. (2013). Developing shared situational awareness for emergency management. *Safety Science, 55*, 1-9.

Slavkovikj, V., Verstockt, S., Van Hoecke, S., & Van de Walle, R. (2014). Review of wildfire detection using social media. *Fire Safety Journal, 68*, 109–118.

Smith, E., & Senter, R. (1967). Automated readability index. *Aerospace Medical Research Laboratories (U.S.)*, 1-14.

Sonnenwald, D. H., & Pierce, L. G. (2000). Information behavior in dynamic group work contexts: interwoven situational awareness, dense social networks and contested collaboration in command and control. *Information Processing & Management, 36*, 461-479.

St Denis, L. A., Palen, L., & Anderson, K. M. (2014). Mastering social media: an analysis of Jefferson County's communications during the 2013 Colorado Floods. In *Proceedings of the Information Systems for Crisis Response and Management Conference (ISCRAM 20014)* (pp. 737-746).

Stanford NLP Group. (2017). Software. https://nlp.stanford.edu/software/ Accessed June 2017.

Stanton, N. A., Chambers, P. R. G., & Piggott, J. (2001). Situational awareness and safety. *Safety Science, 39*, 189-204.

Temnikova, I., Vieweg, S., & Castillo, C. (2015). The Case for Readability of Crisis Communications in Social Media. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 1245–1250): ACM.

Treem, J. W., & Leonardi, P. M. (2013). Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association. *Annals of the International Communication Association, 36*, 143-189.

van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS Quarterly, 28*, 695-704.

Verma, R., Jahn, B., & Kunz, W. (2012). How to transform consumers into fans of your brand. *Journal of Service Management, 23*, 344-361.

Verma, S., Vieweg, S., Corvey, W. J., Palen, L., Martin, J. H., Palmer, M., Schram, A., & Anderson, K. M. (2011). Natural Language Processing to the Rescue? Extracting" Situational Awareness" Tweets During Mass Emergency. In *ICWSM*: Citeseer.

Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1079–1088): ACM.

Watson, H., & Rodrigues, R. (2017). Bringing Privacy into the Fold: Considerations for the Use of Social Media in Crisis Management. *Journal of Contingencies and Crisis Management*, n/a-n/a.

Yang, C., Chen, J., & Su, G. (2016). Effective situational awareness to wildfire emergency command based on multi-model forecasting system. In *Proceedings of the Second ACM SIGSPATIAL International Workshop on the Use of GIS in Emergency Management*. Burlingame, CA: ACM.

Yang, Y., & Liu, X. (1999). A re-examination of text categorization methods. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 42-49). Berkeley, CA, August 15-19: ACM.

Yates, D., & Paquette, S. (2011). Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. *International Journal of Information Management, 31*, 6-13.

Yin, J., Lampert, A., Cameron, M., Robinson, B., & Power, R. (2012). Using social media to enhance emergency situation awareness. *IEEE Intelligent Systems*, *27*, 52-59.

Yogatama, D., Dyer, C., Ling, W., & Blunsom, P. (2017). Generative and discriminative text classification with recurrent neural networks. *arXiv preprint arXiv:1703.01898*.

Yu, B., & Kwok, L. (2011). Classifying business marketing messages on Facebook. In *Association for Computing Machinery Special Interest Group on Information Retrieval*.