Onorati, T., Díaz, P., Carrión, B. (2019). From social networks to emergency operation centers: A semantic visualization approach. *Future Generation Computer Systems*, 95, pp. 829-840.

DOI: 10.1016/j.future.2018.01.052

From Social Networks to Emergency Operation Centers: a Semantic Visualization Approach

Teresa Onorati[a,*], Paloma Díaz[a], Belen Carrion[a]

[a] *Computer Science Department, University Carlos III of Madrid Avda de la Universidad 30, 28911 Leganes, Madrid, Spain*

Abstract

Social networks are commonly used by citizens as a communication channel for sharing their messages about a crisis situation and by emergency operation centers as a source of information for improving their situation awareness. However, to utilize this source of information, emergency operators and decision makers have to deal with large and unstructured data, the content, reliability, quality, and relevance of which may vary greatly. In this paper, to address this challenge, we propose a visual analytics solution that filters and visualizes relevant information extracted from Twitter. The tool offers multiple visualizations to provide emergency operators with different points of view for exploring the data in order to gain a better understanding of the situation and take informed courses of action. We analyzed the scope of the problem through an exploratory study in which 20 practitioners answered questions about the integration of social networks in the emergency management process. This study inspired the design of a visualization tool, which was evaluated in a controlled experiment to assess its effectiveness for exploring spatial and temporal data. During the experiment, we asked 12 participants to perform 5 tasks related to data exploration and fill a questionnaire about their experience using the tool. One of the most interesting results obtained from the evaluation concerns the effectiveness of combining several visualization techniques to support different strategies for

*Corresponding author

Email addresses: `tonorati@inf.uc3m.es` (Teresa Onorati), `pdp@inf.uc3m.es` (Paloma Díaz), `bcarrion@pa.uc3m.es` (Belen Carrion)

solving a problem and making decisions.

---

1. Introduction

Social networks, such as Twitter, Facebook, and Instagram, have changed the means by which people communicate with anyone from anywhere at any time. The publication and sharing of opinions, feelings, or experiences is becoming a new paradigm for being aware of occurrences in ones surroundings. This is particularly true when some exceptional circumstances occur. For instance, the first news about the terrorist attacks of November 2015 in Paris originated from a tweet sent from the stadium reporting an explosion. Then, millions of tweets were published by people sharing a personal situation, praying for the victims, or following the response activities. Among the shared messages, it is worth noting spontaneous initiatives, such as the hashtag #porteouverte (open door, in English), where a citizen originated a self-initiated plan that offered people a safe place where they could stay until the crisis was concluded. Emergency operators could be interested in identifying these kinds of spontaneous actions and include them in their official response plan, integrating citizens acting as *first-first-responders* [1, 2, 3].

The growing usage of Twitter represents a broad source of information for agencies and organizations involved in crisis management, such as news channels or emergency operation centers [4, 2, 3, 5, 6, 7]. However, to integrate the tweets sent during a crisis into a common operational picture, emergency operators and decision makers have to deal with a large set of unstructured information, the content, reliability, quality, and relevance of which may vary greatly [3, 8, 9]. The filtering of the resulting dataset to obtain a manageable knowledge source can be a time-consuming task and a barrier to guaranteeing the efficient integration of social network usage in emergency operation centers [2, 3].

A visual analytics solution can meet this challenge by combining analytical reasoning techniques and information visualizations to improve the management of large datasets. Examples already exist of visual analytics tools developed to deal with social networks that focus on the processing and visualization of a specific context of use and set of information, such as topic identification [10, 11, 12], event detection [13, 11, 14], and sentiment analysis [15]. This means that, to adopt a visual analytics approach, researchers need to determine the most effective techniques for the context of emergency management (EM) to analyze and visualize information originating from social networks. It is crucial to design visualizations that are easy to interpret and provide efficient answers to the questions of operators and decision makers.

To design an appropriate solution, an understanding of which information from social networks can facilitate crisis management, how it should be presented, and in which phase it is to be used is essential. To answer these questions, we recruited 20 practitioners to participate in an exploratory study. The results show that they agreed that three types of functionalities are crucial for obtaining knowledge about citizens' generated content: geolocation, topic identification, and topic search. Our proposed semantic visualization tool addresses these three requirements by integrating information collected from Twitter to support sensemaking in the decision-making process of EM operators. Content shared on Twitter is collected and analyzed semantically, focusing on its relevance to the specific context, by using a semantic modeling technique already presented in [16]. Filtered information is then displayed in various visualization forms combining several dimensions, such as geographic maps, topic relevance, or data frequency, each one of which was designed to represent a different point of view for interacting with the same dataset.

As a first step in the deployment of the visualization tool, we conducted an experiment to assess its usability and test whether it can support typical information exploration tasks. For this purpose, we asked participants to perform several tasks taken from the categories proposed in Andrienko and Andrienko's paper [17] to explore spatiotemporal data. In particular, we focused on those related to the analysis of individual data elements, such as the most frequent term or the highest number of published tweets. The participants in this experiment were 12 citizens whose profile was different from that of the experts conducting the exploratory study, because at this stage the goal was to assess not the tools potential acceptance by real operators but rather its capacity to support the exploration of large and unstructured datasets. The evaluation confirmed that it is quite easy to understand and interact with the visualizations. In particular, we noticed that different participants utilized different visualizations to solve the same problem, suggesting that the provision of multiple visualizations of the same dataset allows users to find that which fits better their manner of exploring data. Thus, the visualization tool provides flexibility in data exploration to meet not only the three aforementioned requirements but possibly additional requirements not identified in the exploratory study.

The rest of the paper first presents in Section 2 a brief introduction of social networks in crisis situations and existing visual analytics tools for managing large datasets. Section 3 describes the exploratory study and the findings that guided the design process. The semantic visualization tool is described in Section 4 and evaluated in Section 5 with details about participants, performed tasks, and results. Finally, some conclusions are drawn and future research in this context is suggested in Section 6.

2. Related Work

2.1. Social Networks in Crisis Situations

In crisis situations, citizens play the crucial role of first-first-responders who start activities and initiatives to help their communities [1, 2, 3]. They are the first on the scene, and therefore, have access to first-hand information about the situation, and they also have a certain capacity to react before the official first responders assume control of the situation. One of the objectives of crisis informatics is to study the manner in which this information can be used to establish an efficient communication channel among citizens, volunteers, and EM workers [18]. In this direction, several information communication technologies (ICTs) have been used to support people to exchange data in the form of blogs and other online platforms [5, 6, 19, 20, 21].

Citizen participation in the EM process has gained relevance with the advent of social networks, such as Twitter, Facebook, and Flickr. The behavior of social networks changes during large-scale events: the number of published messages increases, as well as the number of users joining them for the first time [7, 22]. In particular, people utilize them to monitor the situation, support charitable causes, or participate in the response process [23]. The information shared on social networks as a direct communication channel among EM operators, victims, and witnesses of the evolution of the crisis or to ask for help can be useful for emergency operation centers [24]. Several surveys conducted using EM practitioners have confirmed the need to integrate social networks into crisis management and to support their daily duties [2, 3, 8]. To allow this, it is crucial to define solutions that facilitate the collection and identification of reliable and useful data from the real-time stream of social messages in order to avoid issues such as lack of trust, overwhelming volume of data, and low quality of data [25, 26, 8]. Algorithms have been developed to meet the challenges involved in evaluating the quality of citizen-generated content [9] or removing noisy redundant data [27]. Since the focus of this research was on information visualization, this section reviews examples of visual analytics tools built to analyze and visualize information collected from social networks.

2.2. Visual Analytics Tools for Social Networks

In this paper, we propose a visual analytics tool as a solution for collecting, analyzing, and presenting information collected from social networks to improve the decision-making process [28]. By combining analytical reasoning techniques with interactive visualizations, this tool enhances the problem-solving capabilities of the decision makers without overwhelming them with data [29].

An interesting contribution in this direction is that of Savikhin et al.s study of the effect of visualization for making decisions over economic data [30]. The authors recruited two groups of undergraduate students with little experience in economic analysis and asked them to draw conclusions after observing a dataset. While the first group used a spreadsheet for this task, the second interacted with a visual analytics tool. Their experimental results showed that the decisions of the second group were more accurate and closer to the optimal solution.

Using a domain-oriented approach, Tomaszewski et al. discussed the manner in which geovisual analytics can support crisis management [31]. This emerging discipline originated in the application of visual analytics over geographic information. The results led to the development of tools for supporting analysts reasoning over geographic data to find knowledge about a specific problem. In particular, the authors reviewed the different phases of the crisis management process to identify the geovisual analytics tasks that enhance the overall situation awareness and decision-making processes, such as dynamic collaboration, data navigation, and performance evaluation.

Among existing visual analytics tools for managing Twitter data, VisAware was developed using a journalistic approach, organizing the information as answers to the traditional questions of *what*, *when*, and *where* [13]. For this purpose, the authors introduced a visualization that highlights existent correlations among available datasets. The tool was tested in two different applications: VisAlert, to represent information about emergency alerts to help operation centers reach decisions about a crisis, and BioWatch, which facilitates the identification and collection of data about harmful agents. Following a similar approach, TweetXplorer offers several visualizations to answer the questions of *who*, *what*, *when*, and *where* [10]. In particular, it includes a network of nodes representing the users publishing tweets (*who*), a word cloud with the most common terms (*what*), a heat map showing the volume of tweets in geographical locations (*where*), and a calendar with different colors according to the number of tweets published each day (*when*).

Twitscoop [11] is a visual tool for exploring trending topics on Twitter represented in a dynamic treemap that changes over time to reflect their real-time evolution in order to identify the events that are currently drawing users' attention. The evolution of Twitter content over time is also one of the most salient features of TwitterScope [32]. This tool monitors a stream of tweets and, for each one, performs a semantic analysis to determine which topic is being discussed. Tweets about the same topic are grouped into clusters and represented as countries in a map metaphor.

Topic clustering is also used in TwitterReporter to discover breaking news [12]. The tool monitors newly published tweets and displays them on a map, if the geographic coordinates are available. Once added to the map, the tweets are grouped geographically into clusters of nearby locations labeled with the most frequent term of the collection. Thus, it is possible to filter the clusters specifying the terms and the temporal frame to be shown.

Focusing on opinion mining, Zimmerman et al. proposed EmotionVis, which performs sentiment analysis of tweets retrieved during a large scale event, such as a sports competition [15]. The results are displayed on a dashboard using different visualizations to allow analysts to see the same data from different perspectives.

A different approach, called ScatterBlogs, was proposed for detecting anomalies in social network content, such as critical and unexpected situations [14]. Based on the assumption that, if an exceptional event occurs, the information is quickly spread using different communication channels, the algorithm compares the most discussed topics collected from different social networks and allows users to change the relevance level of the topics according to their interests.

Table 1 summarizes the details of some visual analytics tools for handling large datasets collected from Twitter, focusing on the type of analysis they perform, the data sources used, and the visualizations proposed. These tools introduce interesting approaches for coping with visual analytics, which were taken into account in our design solution. Thus, one of the proposed mechanisms is based on the frequency of the terms used in the social content, filtering those that have received the greatest attention [10, 11, 12] so that the most common words are also considered the most relevant ones. Another interesting data analysis addresses event detection, allowing exceptional circumstances currently occurring to be identified [13, 11, 14]. This analysis can help users be aware of the latest news even before it is published on the official channels.

To handle the volume of information generated on social networks about a topic, the reviewed tools offer different analyses, including the evolution of the number of tweets along time [10, 11, 14], categorization into clusters based on the semantic similarity [11, 32, 12], sentiment analysis [15], and geographical location [10].

Depending on the offered data analysis, visual analytics tools also provide different visualization techniques with which the user can interact, including statistical graphs, dynamic treemaps or heat maps for volume, maps for clustering tweets, word clouds for term frequency, and concentric rings for event detection. In this study, we combined the data analysis and visualization approaches implemented in these tools to build a flexible tool for data exploration, where the term flexible means that the tool provides several means of exploring data and interacting with the visualizations. It is noteworthy that no unique visual representation for data analysis exists that can satisfy the needs of all decision makers at all stages of a crisis. The choice of visualization depends not only on the questions for which the user requires an answer and the data that he/she needs to be displayed, but also on the strategies he/she follows to understand the problem and search for solutions. For this reason, we posit that visualization tools need to provide a variety of options to fit different sensemaking procedures.

Table 1: Summary of visualization tools

Tool [ref]	Data analysis	Data source	Visualization
VisAware [13]	event detection	emergency alerts	concentric rings and nodes
TweetXplorer [10]	term frequency, volume of tweets, geographical position	tweets about emergency	users network, word cloud, heat map, calendar
Twitscoop [11]	term relevance, evolution of tweets volume, synonym-based categorization, event detection	trending topics	dynamic treemap
TwitterScope [32]	synonym-based categorization	Twitter	map of clusters
TwitterReporter [12]	synonym-based categorization, term frequency	breaking news	map
EmotionVis [15]	sentiment analysis	Twitter	dashboard
ScatterBlogs [14]	event detection, volume evolution	Twitter	map, word cloud

3. Exploratory Study

The literature analysis confirmed that it is necessary to support and guide emergency operators in integrating the use of social networks into their daily practice. In this paper, we propose using a visual analytics solution to meet this challenge. We noticed that several types of data analysis and visualization techniques are implemented in the various examples of visual analytics tools that can be used to interact with the information collected from social networks. In particular, the choice of a specific approach for analyzing and visualizing data depends on the context in which it is expected that the tool will be used. This context includes the users, the tasks to be performed, and the conditions under which the tasks are performed. In the specific context of integrating social network data in the crisis management process, we identified three open questions: (i) What information should be extracted from social networks, that is, which data can be relevant for decision makers? (ii) How should information be presented, that is, which are the best visualizations to support decision making? and (iii) When, that is, in which phase of the EM process, can visualizations be used? To find answers to these questions, we conducted an exploratory study involving several experts for framing the context of this study [33]. We recruited 20 practitioners who were attending a course on the use of information technology to improve cooperation during crisis management organized by the Spanish School of Civil Protection. After a lecture on citizen participation and the use of social networks, they volunteered to take part in our study and answer a paper-based questionnaire. The items in the first part addressed ethnographic information about their expertise in the area. As shown in Figure 1, the responses showed that the majority (90%) had more than five years' experience in operative procedures, coordination, supervision, or telecommunication infrastructures (see Figure 2). Their areas of expertise varied among different types of crisis, including natural disasters, environmental, health, domestic, and traffic accidents, meteorological events, and rescues. In light of this profile, it was considered that the group could potentially provide varied and valid insights into the integration of social networks in the EM practice.

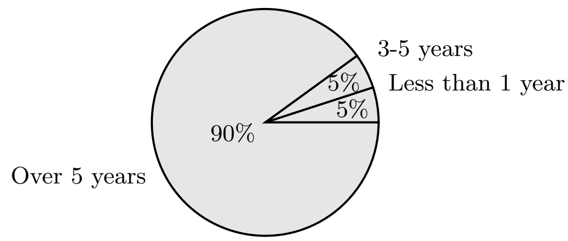

Figure 1: Years of experience of participants in the exploratory study

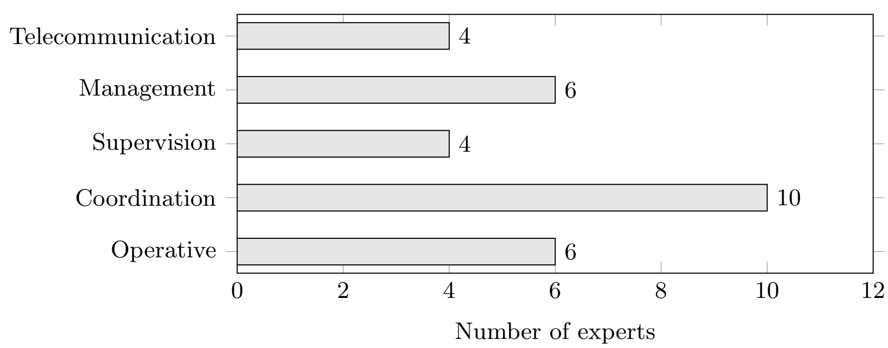

Figure 2: Roles played in the emergency management process by participants in the exploratory study

The following questions explored the *what*, *how*, and *when* issues. All questions had multiple-selection closed-ended options derived from similar studies, in particular those already described in [3], and a final open-ended "other" field for providing additional feedback. In the first question, participants were asked to identify the benefits they could gain from utilizing social networks, including items such as improvements in communication with citizens, volunteers, and other agencies, improvement in their own understanding of the situation, improvement in their own understanding of the social effect of the situation, and the ability to anticipate potential risk situations. In their responses to the second question, they were asked to indicate the type of data that would be more valuable for them, including common topics or hashtags, the location of tweets, geographical areas with more or less activity, and tweets about a specific topic. To reinforce the validity of this opinion, they were asked in the third question about the potential uses of information, such as to improve communication, detect needs and problems, activate groups of volunteers, or identify citizen initiatives. Finally, a fourth question asked about the specific phases of the EM process (namely, early warning, preparation, response, and recovery) in which information extracted from social networks could be useful. At the end of the survey questionnaire, we presented an open-ended question about any other potential usages not included or explored in the previous questions.

Table 2 shows the results of the closed-ended question options. In their responses to the final open-ended question, the participants suggested interesting scenarios, such as "*To let us know where and what is happening for improving the response*." An example is sending notifications in case of incidents, "*to alert citizens about accidents on the roads and prevent traffic jams*," or establishing a bidirectional channel "*to guarantee an effective communication with rural or sparsely populated areas*."

Table 2: Results for the closed-ended questions in the exploratory study

Topic	%
Benefits of the social networks' usage	
Improves communication with citizens, agencies, or volunteers	65%
Improves the understanding of what is happening in a geographic area	45%
Improves the understanding of the social impact of an event	40%
Anticipates potential risks	50%
Information to be retrieved from social networks	
Most discussed topics	25%
Location of users communicating about a specific topic	45%
Areas with higher or lower activity in social networks	20%
Published messages about a specific topic	45%
Practical examples of the social networks' usage	
To improve communication with citizens, agencies, or volunteers	45%
To detect citizens' needs or problems	80%
To activate groups of volunteers	30%
To identify citizens' initiatives being broadcast on social networks	25%
Phase of the emergency management	
Early Warning	60%
Pre-emergency or prevention	45%
Response	70%
Recovery	40%

3.1. Findings of the Exploratory Study

According to the comments collected in the exploratory study, the integration of social networks in an EM operation center depends on the decisions that need to be made, the tasks to be performed, and the situation to be managed. In the questionnaire analysis results (see Table 2), we can observe some interesting findings for answering the open issues identified by the literature analysis.

Concerning the information that should be filtered from social networks (see the second topic in Table 2), the practitioners indicated three main items (see Table 3): (R1) the geographical location of the social activities (messages/citizens), which can also be used to support the fourth item in the table, which received fewer votes, but can still be relevant (for instance, not having network activity in an affected area might indicate a connectivity problem); (R2) the most relevant topics; and (R3) messages about a specific event. It is worth noting that the practitioners answers indicate no clear threshold among the four options, making them all relevant requirements. Moreover, they also described a wide range of activities that would benefit from using these data, such as the communication among all the actors involved in the crisis management, discovering what is happening in a specific geographic location, and understanding the effect that a crisis has on the population.

Table 3: Requirements in the exploratory study

ID	Requirements
R1	Geolocation
R2	Topic Identification
R3	Topic Search

Another interesting fact that emerged from the exploratory study concerns the management phase, during which the social networks would be more helpful for the operation center. Although all the participants considered that social network information could be useful in all four phases, the early warning and response phases obtained the highest scores. These two phases are completely different and are performed under very different conditions: whereas the early warning phase constitutes the process of monitoring a situation to detect whether it could evolve into a crisis, the response phase involves solving an existing crisis, minimizing damages, and returning conditions to normal as soon as possible. The early warning phase is not usually executed under stressful circumstances and might benefit from integrating citizens in EM to collect information, following a paradigm similar to that applied in citizen science, as discussed in [3]. During the response phase, once the crisis has occurred, real-time streaming of messages guarantees that operators can establish an effective communication channel with the citizens to guide their participation and collect witnesses, in particular during the first moments after the crisis occurs when no official help is yet present in the area. Thus, it is possible to collect fresh and updated information to be used to assist better decision making. The participants stressed that this is a real use in rural and isolated communities, but studies in the literature confirmed that it also effective in populated areas for mobilizing the social capital of those affected by the crisis [2, 3]

In conclusion, practitioners would value a tool for searching and filtering social messages and accessing three main data groups: geolocation, topic identification, and topic search. Moreover, they would like a solution that offers various points of view of the same dataset that can be easily used as a support for different activities, such as monitoring an event (real-time) or performing a post-event analysis.

These results led us to a significant research question: **How can social networks be filtered and visualized to meet the practitioners' requirements?**. As shown in the next section, to address this question we propose a visual analytics solution in which a semantic analysis mechanism and several visualization techniques are combined. Indeed, the visualizations would support the EM operators in navigating the information they need and achieving their goals using different points of view.

4. Semantic Visualization Tool

The semantic visualization tool proposed in this paper is based mainly on the findings of the exploratory study and the analysis of similar contributions in the literature. To design a system that allows users to interact with the information extracted from the social networks, we followed a *details-on-demand* approach that offers the possibility to visualize the big picture of a current situation, as well as the individual elements that depend on the context and the objectives

[34].

The architecture of the tool comprises two different levels (see Figure 3): information recovery and information visualization. During the information recovery phase, following the requirements listed in Table 3, a semantic analysis module categorizes the collected data to facilitate searching for and filtering a specific topic. For information visualization, the tool offers five complementary representations of the same dataset (treemap, word cloud, bubble chart, animated map, and infowindow) that EM operators can use according to the information that they are searching and the manner in which they interpret and reason over the data. As shown below, they can also interact with the visualizations, using filters that facilitate their exploration. The following subsections describe further both levels: information recovery and information visualization. In these descriptions, the term user refers to the potential user of our visual analytics solution.

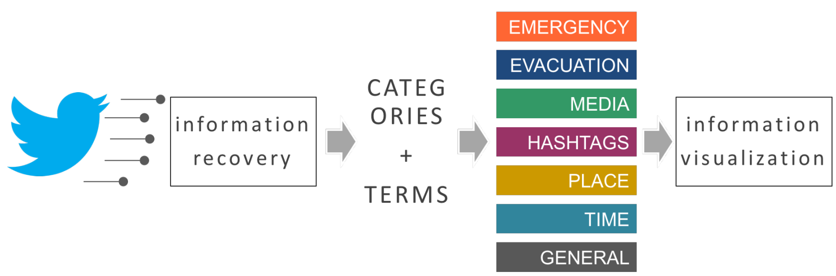

Figure 3: Architecture of the semantic visualization tool

4.1. First Level: Information Recovery

The first level of the tools architecture represents the underlying logic and utilizes a method defined in a previous paper [16] to analyze a dataset of tweets semantically. The first step of the method consists of collecting the tweets using Twitter's streaming application programming interface (API) [35]. This API provides developers with access to the stream of messages shared on Twitter, allowing real-time data recovery, which can be restricted to certain users or topics. Our tool uses keywords chosen by the user to retrieve and store the tweets containing them (*Topic Search* - requirement R3 in Table 3). Users can then look for any word in which they are interested, including places, nouns, verbs, or hashtags. When the keywords have been entered in the search box, a real-time process starts constantly retrieving new posts related to the keywords. Data recovery can be disabled (and re-enabled) by the user in cases where, for example, a term is relevant only for a limited period or the number of publications increases massively and becomes unmanageable.

The collected tweets are semantically analyzed using the Stanford Part-of-Speech Tagger [36], which assigns the corresponding part of speech (i.e., its syntactic function) to each word. This is done for the purpose of extracting nouns from the tweets and reducing them to their root (i.e., with no prefix or suffix). Groups of nouns that appear together are taken as single terms, except for usernames and hashtags, such as *Eiffel Tower*, *President Hollande*, or *concert hall*.

To identify the main topics in the dataset (*Topic Identification* - requirement R2 in Table 3), the obtained list of terms is semantically related to a category using specific knowledge models (namely, an ontology and two taxonomies) and semantic relations extracted from a dictionary, i.e., WordNet. Considering the application domain, we defined eight categories as answers to six journalistic questions [16]: "Who shared the information (*username* and *media*)? "What happened? (*hashtags*, *emergency*, and *evacuation*); "Why did it happen? (*emergency*); "Where did it take place? (*place*) "When did it take place? (*time*); and "How did it happen? (*emergency*). In particular, the category *username* contains all the terms that begin with the "at sign (@), as in Twitter it is used to mention a user. Similarly, the category *hashtag* contains all the terms that begin with a hash sign (#). Finally, the category *general* contains all the terms that do not fit into any of the other existing categories. Except for the last three categories, the interface provides users with tools to modify the information visualization level.

According to the search term chosen by the user, the volume of collected information can increase dramatically. To guarantee an effective and efficient solution, the semantic analysis is performed on a separate server and integrated as a Web service. Thus, even if the data recovery is disabled, the database storing all the data is continuously updated. Moreover, before saving a tweet, an evaluation of its content, leaving out retweets, emoticons, or messages in other languages, is performed.

4.2. Second Level: Information Visualization

The second level of the architecture consists of an interactive interface organized into three columns (see Figure 4). The menu on the left hand side answers the third requirement discovered by the item in the exploratory study questionnaire that addressed topic search (see Table 3). It provides a search box for querying, a list of the active queries, and the tweets related to the queries. The central pane displays the current visualizations with a timeline at the bottom representing the evolution of the dataset. Users can interact with the timeline, going back and forward to visualize the evolution of the tweets published in a certain time interval. In the pane on the right hand side, users are provided with tools for interacting with the visualizations, whether to change the level of detail or change the type of visualization. In particular, a list of included terms (pink bar in the figure) with the categories (blue bar in the figure), specific filtering mechanisms (green bar in the figure), and the set of available visualizations (yellow bar in the figure) are provided.

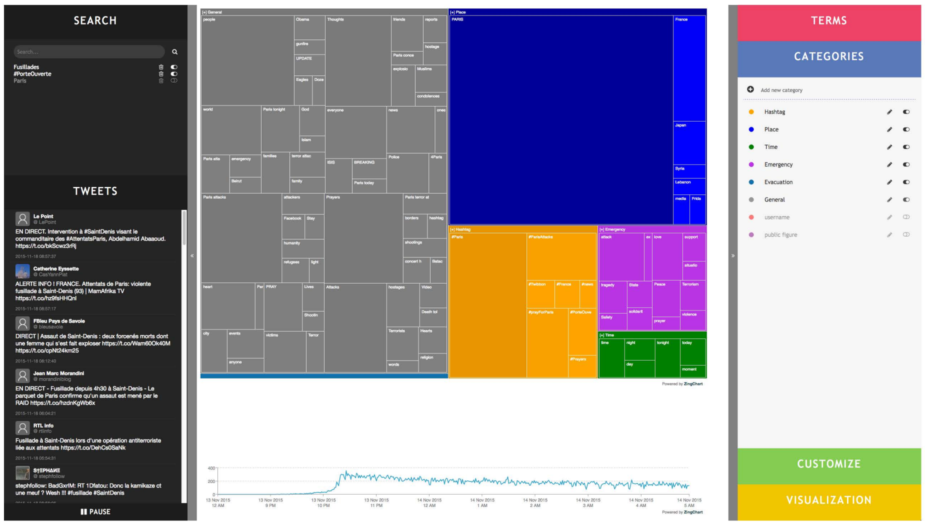

Figure 4: Main interface of the semantic visualization tool

Visualization plays a fundamental role in facilitating data interpretation and consequently in improving the decision-making process [28]. As stated by the EM experts in our exploratory study, the provision of different points of view to navigate the same dataset and to support different activities is crucial. Bearing this idea in mind, we designed the tool such that it offers five different visualizations: the *treemap*, the *word cloud*, the *bubble chart*, the *animated map*, and the *infowindow* (see Figure 4), each of which is focused on one of the specific requirements derived from the exploratory study. Moreover, to guarantee a detail-oriented view, we bore in mind Shneiderman's visual information-seeking mantra [34]: *overview first, zoom and filter, details on demand*. Thus, users can interact with the visualization, changing the zoom level and using filtering mechanisms if they need more specific information.

The first technique is the ***treemap***, in which data are represented using nested rectangles, where the child and parent items have a hierarchical relation. This technique was introduced by Shneiderman to organize the file system of a computer [37], taking advantage of the existent hierarchical relation between folders and files. Following the same idea, we used the treemap to represent the second requirement concerning topic identification derived from the exploratory study (see Table 3). Considering the results from the information recovery level, the hierarchical relationship is shown between the terms extracted from social messages and the main topics given by the eight semantic categories, i.e., emergency, evacuation, media, username, hashtags, time, and place. Each term is represented as a rectangle, the color of which depends on the related category and the size of which is determined by the frequency of the term. Figure 5 shows a detail of the treemap for two categories: *hashtags* and *emergency*. This visualization can also facilitate identification of patterns over the collected information, such as the most relevant topics (e.g., the #Paris box is larger and darker than the #news box in Figure 5), and the most interesting categories (e.g., emergency in Figure 5 is the most populated category with 13 boxes).

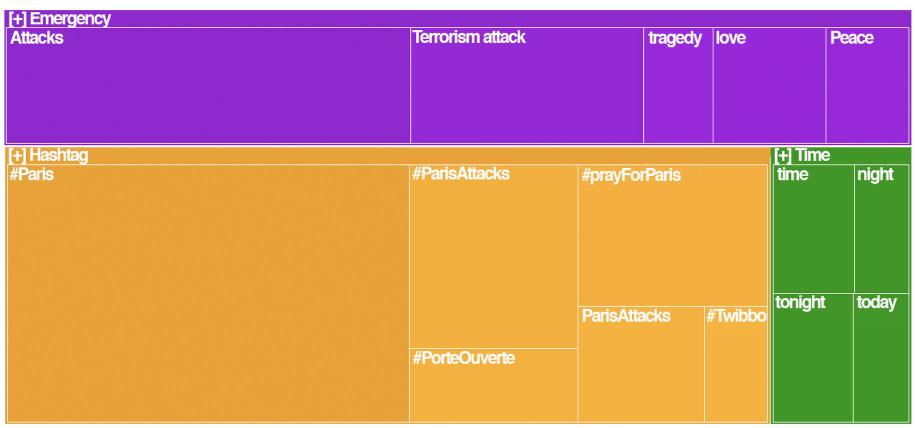

Figure 5: Part of the treemap visualization showing the categories emergency (purple boxes) and hashtag (orange boxes)

The second technique is the ***word cloud***. Born outside the computer science area for mapping mental reasoning processes, this technique became well-known for representing large lists of words and their frequencies [38]. For this reason, we chose it as an alternative answer to the second requirement about topic identification (see Table 3). The frequency of each term is represented by a proportional font size, while the font color changes according to the topic to which the term is related. The difference between this visualization and the treemap visualization is in the organization of the terms in a circle, where the central terms are the most relevant. Thus, the word-cloud visualization can be particularly useful for detecting real-time trending topics. In fact, the topics that are presented in a larger font size are those that have received more mentions in citizens messages and are consequently the most relevant for them. Moreover, the most common terms appear in a central position, thus facilitating their detection. Figure 6 shows an example of the words collected during the Paris attacks. In the representation, we can observe that the font of the term "attacks" is larger and, consequently, more relevant than the term "reports."

Figure 6: Word cloud visualization

The ***bubble chart*** is one of the design visualization techniques for linear data that are easiest to understand [39]. To take more advantage of the space, this chart uses labeled circles of different sizes according to the data represented. Because of its characteristics, we decided to provide users with a bubble chart as the third option for the topic identification requirement (see Table 3). The terms are represented by circles (or bubbles) grouped according to topics identified by colors (see Figure 7). The size of each circle is determined by the term frequency. For example, in Figure 7 the bubble for the term "Prayers" is larger than that for Japan. To offer a more detailed view, the bubbles are presented in a twodimensional space, where the y-edge is the term frequency and the x-edge is the topic. This visualization combines the advantages of the treemap (i.e., pattern identification) and the world cloud (i.e., its topic-oriented nature). Additionally, the number of included terms is lower than in the previous visualizations, and this can offer users a better navigation experience. Moreover, it can be useful for small-scale events, such as traffic accidents or crises in rural areas.

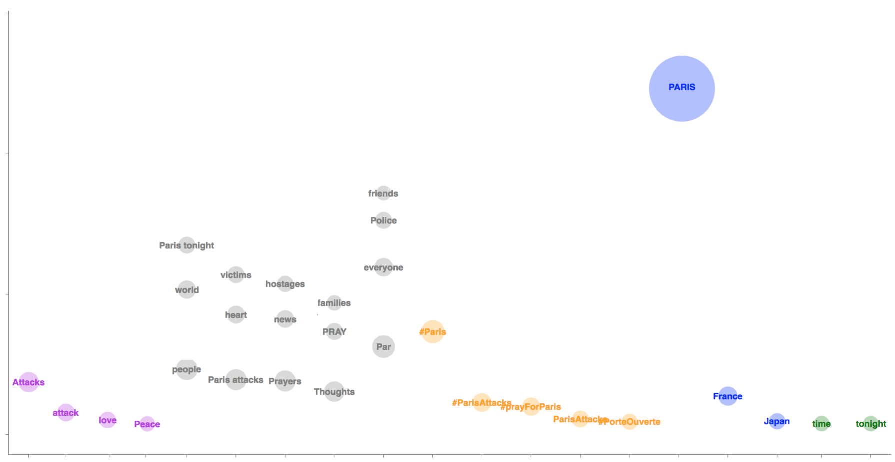

Figure 7: Bubble chart visualization

The fourth visualization is the ***animated map***, a map with clusters of geographically close tweets (see Figure 8). The usage of this technique to represent real-time streaming of social messages in the context of crisis management has already been presented in [40, 41]. The tweets are located sequentially depending on their freshness, that is, their publication time, and their appearance is animated as a raindrop. Thus, users can observe the evolution of the volume of tweets and relevant topics over time with respect to their location, as well as to the location from which people are sending the messages. These activities respond to the first requirement about geolocation derived from the exploratory study (see Figure 7).

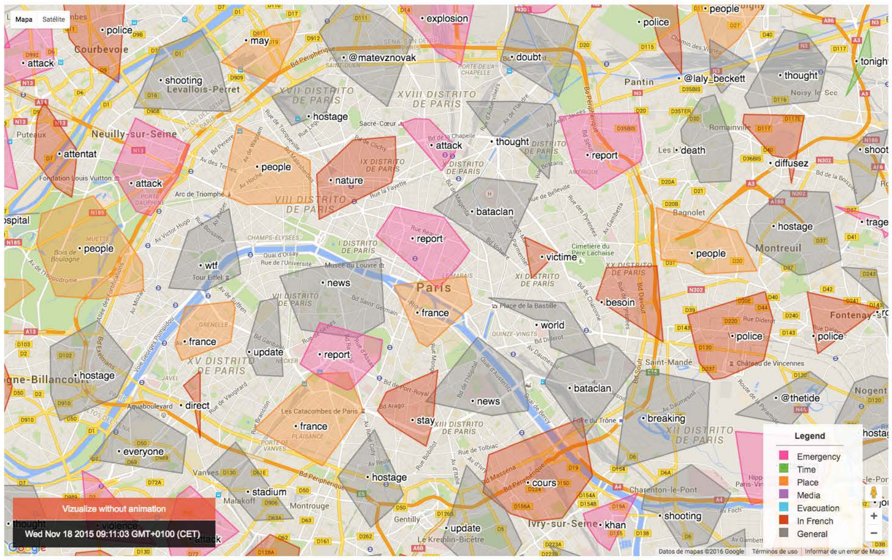

Figure 8: Map visualization

When tweets are located on the map, they are added to the closest cluster. Clusters are created according to the geographical proximity of the tweets to form a convex hull (i.e., the smallest convex polygon that can contain them) using the Graham scan technique [42]. The maximum geographical distance for determining whether a tweet should be included in a cluster was determined experimentally, and it changes with the zoom-in and zoom-out navigation. The minimum zoom level shows no clusters and tweets are located on the map individually with a marker. To view the content of a specific tweet, users can click on a marker that opens an infowindow showing details such as the author, geographical position, and the text of the message (see Figure 9).

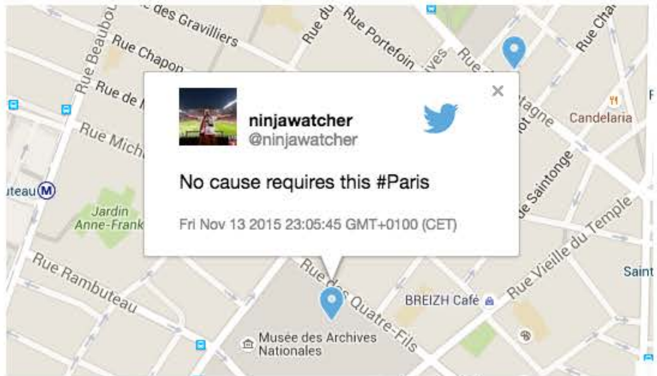

Figure 9: Infowindow

5. User Evaluation

The evaluation of visualization tools is usually considered a complex task. On the basis of an extensive literature review, Lam et al. identified seven evaluation scenarios and the objectives, questions, and techniques that can be used for each of them [43]. The authors analyzed several contributions from the information visualization research community, attempting to find a relationship between the goals that should be evaluated and the evaluation techniques that should be applied. As a result, seven evaluation scenarios were identified: environments and work practices, visual data analysis and reasoning, visualization and communication, collaborative data analysis, user performance, user experience, and automated evaluation of visualizations. As a first step in the deployment of our proposed tool that may be used by real EM operators, we were interested in understanding whether it can be an effective solution for filtering and visualizing information arriving from social networks to meet the requirements identified during the exploratory study (see Table 3). Whether certain requirements are met by a new tool can be determined by using a controlled experiment in which users are asked to perform elementary exploration tasks [17], for which no specific knowledge in the domain is required. The goal is then to evaluate the users experience, for which purpose laboratory questionnaires are used [43]. The following subsection provides further information about the evaluation procedure and findings.

5.1. Participants

In this evaluation, our objective was to focus mainly on the user experience using the visual analytics tool, focusing in particular on the visualization techniques. For this purpose, we recruited 12 participants with no EM background. As stated previously, the type of tasks to be performed did not require knowledge in the application domain. Moreover, the utilization of users with no expectations that the tool will improve their daily performance avoided moving the focus of the evaluation to different issues related more to the tools acceptance, an evaluation that will be conducted in a second step when the tool has been proven to be potentially usable and useful. Most of the participants had a background in computer science and their age ranged between 18 and 34 years-old. Because the semantic visualization tool uses data retrieved from Twitter, we sought people who actively use social networks (83% of the participants) or are familiar with them (the remaining 17%). In Table 4, additional demographic details are listed.

Table 4: Demographic data for $N = 12$ participants

Items	Category	%
Gender	female	25%
	male	75%
Age	18-24	42%
	25-34	42%
	35-44	8%
	45-54	8%
Occupation	bachelor student	17%
	master student	41%
	PhD student	17%
	software developer	17%
	ICT services manager	8%
Social networks	signed up	83%
	not signed up	17%
Twitter	signed up	83%
	not signed up	17%

5.2. Data

After their demographic data had been collected, the participants received a short introduction to the tool. The dataset used for the experiment was a collection of tweets published from 13th November to 18th November, 2015, when several terrorist attacks occurred in Paris. During this event, we observed particular scenarios in which the citizens used social networks as a communication channel with authorities to request help or offer shelter.

The information retrieval phase in fact consisted of running queries with the following keywords: *Paris*, *fusillade* (shooting, in English), *Paris attacks*, and *#porteouverte* (open doors, in English). The last hashtag was used during the crisis by people to welcome those seeking shelter to their houses. During the information retrieval phase, we collected around 500000 tweets and reduced them to a total of 119186 after filtering those that were not geolocalized, retweets, emoticons, and short messages containing only one word of less than three characters. After the semantic analysis, 234010 terms were extracted and categorized. Examples include the term *peace* in the category *emergency*, with 735 occurrences where people started to show their grief and indignation against the attacks, *#news* in *hashtag* with 325 occurrences and shared mainly by official news channel profiles, and *hostages* in *general* with 960 occurrences.

5.3. Tasks

Because the tool was designed to handle geographical and temporal information, to identify the tasks for the experiment we applied Andrienko and Andrienko's categorization [17]. These authors analyzed current techniques and methods used to explore spatial and temporal data. In the results of their analysis, two main task categories were identified: *elementary tasks* to address individual data elements and *synoptic tasks* to address the general views of the entire dataset. Considering that the visualizations offer a *details-on-demand* navigation of the data, we focused on the *elementary tasks*. This category was additionally split into three types: *lookup* to search for specific data attributes, *comparison* to match different data values, and *relationship search* to find complex relations among the data attributes. Inspired by this categorization and according to the requirements identified in the exploratory study (see Table 3), we defined the tasks that should be performed in the evaluation as follows.

1. Find the most frequent term (lookup task);
2. Identify which terms appear in more than 4000 tweets (lookup task);
3. Compare the most frequent term in Spain with the most frequent term in France (comparison task);
4. Determine whether the highest volume of tweets was reached before or after the lowest (relationship search task);
5. Determine when the term #porteouverte was first used.

As can be seen, these tasks constitute generic exploration activities not requiring any knowledge of EM. The participants were asked to perform the five tasks using the tool. One researcher was responsible for controlling the experiment, taking notes about problems, suggestions, or any other comments, and observing the strategies adopted by the participants to solve the tasks, that is, the visualization techniques they used for each task.

5.4. Results

One of the most interesting results of the observation of the experiments concerns the different visualizations that the participants explored for solving the tasks. In particular, for the first task, finding the most frequent term, different techniques were used by different participants: the word cloud, the list of terms in the yellow bar on the right side of the tool (see Figure 4), the bubble chart, and the treemap. The most complex task was the fifth (i.e., to determine when the term #porteouverte was first used). This task required the participants to use a combination of one of the visualizations and the timeline. When the time frame in the timeline is changed, the data in the visualization are updated. Thus, it is possible to identify the moment at which the hashtag #porteouverte was published for the first time. Participants mentioned two main concerns when performing this task: first, they failed to see the correlation between the timeline and the visualizations in the central pane of the tool (see Figure 4), and second, they needed more details in the visualizations, such as the frequencies or a direct link to the list of tweets containing a specific term.

After the experiment, the participants were asked to complete a questionnaire evaluating their experience with using the tool. This questionnaire consisted of nine statements, which the participants scored on a 5-point Likert scale: strongly agree (SA), agree (A), neither agree nor disagree (N), disagree (D), and strongly disagree (SD). The statements were derived from the Unified Theory of Acceptance and Use of Technology (UTAUT) model [44]. The UTAUT model defines a list of questions for evaluating users experience using a technology, grouped into four constructs: performance expectancy, effort expectancy, social influence, and facilitation conditions. Considering the scope of this study, we selected a subset of questions from the constructs about performance expectancy and effort expectancy. The results of the questionnaire are shown in Table 5 in terms of mean and standard deviation. The Likert scale is represented by numeric values from 1 (SD) to 5 (SA) to calculate these statistical functions.

Table 5: Mean (μ) and standard deviation (σ) for the questionnaire responses

ID	Item	μ	σ
1	Using this system enables me to accomplish tasks more quickly	4.2	0.6
2	Using this system improves my job performance	4.3	0.6
3	Using this system would make it more difficult to do my job	1.7	0.6
4	I would find this system useful in my job	4.3	0.7
5	I would find it easy to make the system do what I need it to do	4.2	0.7
6	Learning to operate the system would be easy for me	3.5	0.7
7	My required interaction with the system is clear and understandable	4.3	0.7
8	It would be easy for me to become skillful at using the system	4.5	0.7
9	I would find the system difficult to use	2.1	0.5

In general, the participants responses showed a high agreement level with almost every one of the statements in the questionnaire. Except for the sixth item, all the means are greater than 4, between “strongly agree and “agree on the Likert scale. For the negative questions (i.e., the third and ninth items), the mean is around 2, corresponding to “disagree. These results show that the tool was considered a quick and useful support for achieving the tasks and improving the participants performance. Moreover, the users found it easy to understand how the system works and which functionalities are provided. The lowest level of agreement was achieved for the sixth item addressing the ease of learning the tool. In this case, the mean is 3.5 and is related mainly to the difficulties the users found while performing the fifth task. However, even if an initial effort is required to become familiar with the tool, with more practice the users can become skillful at interacting with it, as shown by the eighth item (mean 4.5).

Because the design of the tasks was based on the literature review and the findings of the exploratory study, these results suggest that the proposed tool can be considered a possible solution for filtering and visualizing information collected from social networks. Moreover, we recognized an interesting behavior when observing the participants that was related to the flexibility of the tool: different participants chose a different visualization technique to perform the same task. However, an empirical evaluation with real users, that is, EM operators, would be required to confirm its suitability for improving their daily practice. This type of evaluation is very costly since practitioners are not widely available for participating in these exercises. For this reason, evaluations with non-experts are frequently used in crisis informatics [8], because they allow a first glance at the potential utility of a technological tool before involving real users.

6. Conclusions and Future Work

The semantic visualization tool introduced in this work may help EM operators build a general understanding of events on Twitter to improve their analysis and decisions, as well as the extent to which they can monitor citizen activities and identify any unusual behavior or relevant back-channel sources.

This study was aimed at understanding the manner in which social networks can be integrated into emergency practice, supporting citizens in their participation as first-first-responders in a crisis situation [1, 2]. The proposed solution is based on the results of an exploratory study involving 20 EM experts that we conducted in order to understand better which information and functionalities they require for utilizing social network data in their daily practice. The EM experts identified the challenges that they face, such as knowing what is happening in a specific geographic area, understanding the impact of the crisis on the population, and acquiring real-time updated information. We conceptualized these results into three requirements to be addressed in a visual analytics tool: geolocation, topic identification, and topic search. The proposed solution meets these challenges at two different levels: information recovery and information visualization.

In the literature, we identified several studies on handling the large and unstructured data generated by social networks [11, 12, 32, 25, 26, 8]. Following the same direction, the information recovery module of our solution includes an innovative semantic approach for identifying the most meaningful data using knowledge models, but at the same time it allows users to change the terms and the categories they use according to the aspect on which they want to place more emphasis. On the information visualization side, users can interact with five complementary representations, each of which was designed to respond to the experts' requirements derived from the exploratory study. In our evaluation of the user experience using the tool, we observed that different participants chose different techniques to perform the same data exploration task. This behavior can be related to the different means by which they solve problems and reach decisions, although further research is needed to understand how the reasoning process can influence the users desired representation of a dataset and how the representation can influence the users reasoning.

An additional interesting outcome of this work concerns the emergency phase, which would benefit most from social network data. By virtue of the real-time streaming, collected datasets are always updated with fresh information coming from citizens. Thus, operators can establish a dynamic channel for improving communication during the response phase. To allow this, scalable mechanisms for handling the large volume of data and improving the real-time information processing are required. However, visualizations of the evolution of the volume of tweets, the topics, or their geographical locations during a crisis can be interesting for post-incident analysis. In particular, EM operators could use this type of statistics to improve future response activities, such as generating useful hashtags and recommending that people share them to facilitate information collection. This leads to the tool being envisioned as a very flexible means to explore information at different EM stages. The goal has to be to support different means of approaching the exploration of data to answer different types of questions. For that reason, and as compared to other similar tools presented in the literature (see Table 1), our proposed system provides a varied group of visualizations, the exploration of which can be tailored by the users to render them more interpretable and useful according to their needs and preferences.

The results of the user experience evaluation provided us with several ideas for improving the design of the tool before its evaluation by real EM operators. In such an evaluation, EM operators will participate in a scenario-based evaluation aimed at comparing the semantic visualization tool with the timeline of news, resources, and activities that is currently used in operation centers. We expect to focus the evaluation on the utility of the visualization techniques and their effectiveness for the specific context of EM, to identify the affordances and limitations of each visualization technique. The effective design of an experiment using EM experts is crucial for fully understanding the efficiency of the tool and exploiting new research opportunities in this context. However, there are practical issues that frequently reduce the effectiveness of this type of evaluation in the domain of EM, including difficulties in finding an agency willing to use the tool in a real scenario or a sufficient number of available EM operators working in agencies that use social networks to allow statistically relevant data to be collected.

7. Acknowledgements

This work was supported by the project PACE grant funded by the Spanish Ministry of Economy and Competitivity [TIN2016-77690-R].

References

[1] R. A. Stallings, E. Quarantelli, Emergent citizen groups and emergency management, Public Administration Review (1985).

[2] I. Mergel, Social media adoption: Toward a representative, responsive or interactive government?, in: Proceedings of the 15th Annual International Conference on Digital Government Research, ACM, 2014, pp. 163–170.

[3] P. Díaz, J. M. Carroll, I. Aedo, Coproduction as an approach to technology-mediated citizen participation in emergency management, Future Internet 8 (2016) 41.

[4] B. R. Lindsay, Social Media and Disasters: Current Uses, Future Options and Policy Considerations., Journal of Current Issues in Media & Telecommunications 2 (2010) 287–297.

[5] L. Palen, S. B. Liu, Citizen communications in crisis: anticipating a future of ict-supported public participation, in: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, 2007, pp. 727–736.

[6] L. Palen, S. R. Hiltz, S. B. Liu, Online forums supporting grassroots participation in emergency preparedness and response, Communications of the ACM 50 (2007) 54–58.

[7] A. L. Hughes, L. Palen, Twitter adoption and use in mass convergence and emergency events, International Journal of Emergency Management 6 (2009) 248–260.

[8] C. Reuter, T. Ludwig, M.-A. Kaufhold, T. Spielhofer, Emergency services attitudes towards social media: A quantitative and qualitative survey across europe, International Journal of Human-Computer Studies 95 (2016) 96–111.

[9] T. Ludwig, C. Reuter, V. Pipek, Social haystack: Dynamic quality assessment of citizen-generated content during emergencies, ACM Transactions on Computer-Human Interaction (TOCHI) 22 (2015) 17.

[10] F. Morstatter, S. Kumar, H. Liu, R. Maciejewski, Understanding twitter data with tweetxplorer, in: Proceedings of the 19th ACM SIGKDD inter-

national conference on Knowledge discovery and data mining, ACM, 2013, pp. 1482–1485.

[11] S. Bhulai, P. Kampstra, L. Kooiman, G. Koole, M. Deurloo, B. Kok, Trend visualization on twitter: Whats hot and whats not?, in: Proceedings of the First International Conference on Data Analytics, 2012, pp. 43–48.

[12] B. Meyer, K. Bryan, Y. Santos, B. Kim, Twitterreporter: Breaking news detection y visualization through the geo-tagged twitter network., in: Proceedings of the ISCA 26th International Conference on Computers and Their Applications, 2011, pp. 84–89.

[13] Y. Livnat, J. Agutter, S. Moon, S. Foresti, Visual correlation for situational awareness, in: Proceedings of the IEEE Symposium on Information Visualization, IEEE, 2005, pp. 95–102.

[14] J. Chae, D. Thom, H. Bosch, Y. Jang, R. Maciejewski, D. S. Ebert, T. Ertl, Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition, in: Proceedings of the VAST IEEE Conference on Visual Analytics Science and Technology, IEEE, 2012, pp. 143–152.

[15] C. Zimmerman, M.-K. Stein, D. Hardt, C. Danielsen, R. Vatrapu, emotionvis: Designing an emotion text inference tool for visual analytics, in: Proceedings of the 11th International Conference DESRIST Tackling Society's Grand Challenges with Design Science, 2016, pp. 238–244.

[16] T. Onorati, P. Díaz, Giving meaning to tweets in emergency situations: a semantic approach for filtering and visualizing social data, SpringerPlus 5 (2016) 1782.

[17] N. Andrienko, G. Andrienko, Exploratory analysis of spatial and temporal data: a systematic approach, Springer Science & Business Media, 2006.

[18] C. Hagar, Using research to aid the design of a crisis information management course, in: Proceedings of the ALISE Annual Conference SIG

Multicultural, Ethnic & Humanistic Concerns (MEH), Information Seeking and Service Delivery for Communities in Disaster/Crisis, 2006.

[19] O. Okolloh, Ushahidi, or testimony: Web 2.0 tools for crowdsourcing crisis information, Participatory learning and action 59 (2009) 65–70.

[20] M. Careem, C. De Silva, R. De Silva, L. Raschid, S. Weerawarana, Sahana: Overview of a disaster management system, in: Proceedings of the International Conference ICIA on Information and Automation, IEEE, 2006, pp. 361–366.

[21] C. Reuter, A. Marx, V. Pipek, 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 (2012) 1–16.

[22] T. Preis, H. S. Moat, S. R. Bishop, P. Treleaven, H. E. Stanley, Quantifying the digital traces of hurricane sandy on flickr, Scientific reports 3 (2013) 3141.

[23] D. E. Alexander, Social media in disaster risk reduction and crisis management, Science and Engineering Ethics 20 (2014) 717–733.

[24] B. R. Lindsay, Social media and disasters: Current uses, future options, and policy considerations, 2011.

[25] P. Díaz, I. Aedo, S. Herranz, Citizen participation and social technologies: exploring the perspective of emergency organizations, in: Proceedings of the International Conference on Information Systems for Crisis Response and Management in Mediterranean Countries, Springer, 2014, pp. 85–97.

[26] L. Plotnick, S. R. Hiltz, J. A. Kushma, A. H. Tapia, Red tape: Attitudes and issues related to use of social media by us county-level emergency managers., in: Proceedings of the International Conference on Information Systems for Crisis Response and Management, 2015.

[27] M. Moi, T. Friberg, R. Marterer, C. Reuter, T. Ludwig, D. Markham, M. Hewlett, A. Muddiman, Strategy for processing and analyzing social media data streams in emergencies, in: Proceedings of the International Conference on Information and Communication Technologies for Disaster Management, IEEE, 2015, pp. 42–48.

[28] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, G. Melançon, Visual analytics: Definition, process, and challenges, in: Information visualization, Springer, 2008, pp. 154–175.

[29] N. H. Lurie, C. H. Mason, Visual Representation: Implications for Decision Making, Journal of Marketing 71 (2007) 160–177.

[30] A. Savikhin, R. Maciejewski, D. S. Ebert, Applied visual analytics for economic decision-making, in: Proceedings of the VAST IEEE Symposium on Visual Analytics Science and Technology, IEEE, 2008, pp. 107–114.

[31] B. M. Tomaszewski, A. C. Robinson, C. Weaver, M. Stryker, A. M. MacEachren, Geovisual analytics and crisis management, in: Proceedings of the 4th International Conference on Information Systems for Crisis Response and Management, Delft, the Netherlands, 2007, pp. 173–179.

[32] E. R. Gansner, Y. Hu, S. North, W. Didimo, M. Patrignani, Interactive Visualization of Streaming Text Data with Dynamic Maps, Journal of Graph Algorithms and Applications 17 (2013) 515–540.

[33] R. K. Yin, Case study research: Design and methods, Sage publications, 2013.

[34] B. Shneiderman, The eyes have it: a task by data type taxonomy for information visualizations, in: Proceedings of the IEEE Symposium on Visual Languages, IEEE Comput. Soc. Press, 1996, pp. 336–343.

[35] Twitter, https://dev.twitter.com/streaming/overview, 2017.

[36] K. Toutanova, C. D. Manning, Enriching the knowledge sources used in a maximum entropy part-of-speech tagger, in: Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics-Volume 13, Association for Computational Linguistics, 2000, pp. 63–70.

[37] B. Shneiderman, Tree visualization with tree-maps: 2-d space-filling approach, ACM Transactions on graphics (TOG) 11 (1992) 92–99.

[38] F. B. Viégas, M. Wattenberg, Timelines tag clouds and the case for vernacular visualization, interactions 15 (2008) 49–52.

[39] F. B. Viegas, M. Wattenberg, F. Van Ham, J. Kriss, M. McKeon, Manyeyes: a site for visualization at internet scale, IEEE transactions on visualization and computer graphics 13 (2007).

[40] T. Onorati, P. Isenberg, A. Bezerianos, E. Pietriga, P. Diaz, Walltweet: A knowledge ecosystem for supporting situation awareness (2015).

[41] T. Onorati, P. Díaz, B. C. Recio, Tweet me and i'll help you: Mapping tweets for emergency operation centers: The case of paris attacks, in: Collaboration Technologies and Systems (CTS), 2016 International Conference on, IEEE, 2016, pp. 31–34.

[42] R. L. Graham, An efficient algorith for determining the convex hull of a finite planar set, Information processing letters 1 (1972) 132–133.

[43] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, S. Carpendale, Seven guiding scenarios for information visualization evaluation (2011).

[44] V. Venkatesh, X. Zhang, Unified theory of acceptance and use of technology: Us vs. china, Journal of Global Information Technology Management 13 (2010) 5–27.