Smart Technologies for Emergency Response and Disaster Management

Zhi Liu
Waseda University, Japan

Kaoru Ota
Muroran Institute of Technology, Japan

A volume in the Advances in Public Policy and Administration (APPA) Book Series

IGI Global
DISSEMINATOR OF KNOWLEDGE

Published in the United States of America by
IGI Global
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Names: Liu, Zhi, 1986- editor. | Ota, Kaoru, 1984- editor.
Title: Smart technologies for emergency response and disaster management / Zhi Liu and Kaoru Ota, editors.
Description: Hershey, PA : Information Science Reference, [2018]
Identifiers: LCCN 2017006791| ISBN 9781522525752 (hardcover) | ISBN 9781522525769 (ebook)
Subjects: LCSH: Emergency management--Technological innovations. | Emergency management--Information technology.
Classification: LCC HV551.2 .S6185 2018 | DDC 363.340285--dc23 LC record available at https://lccn.loc.gov/2017006791

This book is published in the IGI Global book series Advances in Public Policy and Administration (APPA) (ISSN: 2475-6644; eISSN: 2475-6652)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

For electronic access to this publication, please contact: eresources@igi-global.com.

Chapter 5
Processing Big Data for Emergency Management

Rajendra Akerkar
Western Norway Research Institute, Norway

ABSTRACT

Emergencies are typically complex problems with serious consequences that must be solved in a limited amount of time to reduce any possible damage. Big data analysis leads to more assured decision making and better decisions can mean greater operational efficiencies, cost reductions and reduced risk. In this chapter, we discuss some issues on tackling emergency situation from the perspective of big data processing and management, including our approach for processing social media content. Communications during emergencies are so plentiful that it is necessary to sift through enormous data points to find information that is most useful during a given event. The chapter also presents our ongoing IT-system that processes and analyses social media data to transform the excessive volume of low information content into small volume but rich content that is useful to emergency personnel.

1. INTRODUCTION

During a disaster, life-saving decisions are often made based on the most current information of a situation and past experiences in similar circumstances. While that's a tried-and-true approach, the availability of complex, computer-generated data streams is changing the ball game for emergency service providers. Hence effective management of emergencies and disasters is a global challenge in big data era. A systematic process with principal goal to minimize the negative impact or consequences of emergencies and disasters, thus protecting societal infrastructure, is called effective emergency and disaster management. It is imperative throughout the world to increase knowledge of emergency and disaster management, for the purpose improving responsiveness. All the above aims may be accelerated by big data analysis.

Big data may be characterized as having four dimensions: Data volume, measuring the amount of data available, with typical data sets occupying many terabytes. Data velocity is a measure of the rate of data creation, streaming and aggregation. Data variety is a measure of the richness of data representation – text, images, videos etc. Data value, measures the usefulness of data in making decisions

DOI: 10.4018/978-1-5225-2575-2.ch005

(Akerkar 2013a). Variability, which represents the number of changes in the structure of the data their interpretation, is a newly added characteristic.

The management of such big data is perhaps one of the key challenges to be addressed by informatics. The wide variety of data acquisition sources available in times of emergency creates a need for data integration, aggregation and visualization. Such techniques assist emergency management officials to optimize the decision making procedure. During the outburst of an emergency, the authorities responsible must quickly make decisions. The quality of these decisions depends on the quality of the information available. A key factor in emergency response is situational awareness. An appropriate, accurate assessment of the situation can empower decision-makers during an emergency to make convenient decisions, take suitable actions for the most affective emergency management.

This chapter is divided into six sections. Section 2 presents various kinds of applications of big data in emergency cycle. Essential smart technological research approaches are discussed in section 3. Various research issues, concerning with big data, are elaborated in section 4. Section 5 describes key challenges and steps for processing social media contents. This section is underlining our approach for emergency management utilizing social media data. The chapter concludes in section 6.

2. BIG DATA AND EMERGENCY CYCLE

Big data is the technological paradigm that enables useful analysis of vast quantities of data to be achieved in practice. Big data is the collection of scientific and engineering methods and tools for dealing with such volumes of data, and addresses not merely the storage but also access to and distribution, analysis, and useful presentation of results (such as visualisation of analysis of the data) for huge volumes of data. Big data is becoming a critical part of emergency communication. Emergency communication does not involve only intentional, explicit exchange of messages – for example first responders talking over a voice connection, or an announcement of a text message warning to citizens threatened by an approaching natural disaster. To be more precise, emergency communication also involves the monitoring and understanding of the complete body of public, openly available communication – such as messages and content being publicly exchanged on social media. Thus, individuals may be reporting their condition to loved ones or making specific requests for help, but a complete analysis of all communications can reveal valuable information of general scope, such as a disease outbreak.[1]

Usually, emergency cycle consists of three phases. "Prevention" and "Preparedness" are conducted *before* an emergency occurs in order to eliminate or reduce the probability of an emergency and to build emergency management capacities. "Response" activities provide emergency assistance to save lives, preserve property and protect the environment *during* an emergency. "Recovery" is the process of returning systems to normal levels *after* an emergency. Big data has been used in all phases of the emergency management cycle as shown in the following Table 1.

Open initiatives and new applications for big data constitute a genuine opportunity to provide decision makers with powerful new tools for tracking and predicting hazardous events, protecting vulnerable communities, understanding human factors and targeting where to optimize programs and policies. For several "data deficient" countries and communities accessing big data can increase credibility and value of meteorological forecasts and warnings. Turning big data sets – satellite images, in situ and mobile sensor observations, online user-generated content, environmental data archives, weather and water forecasts, and climate model results, etc. – into useful and actionable information and integrating this complex

Table 1. Data types and various phases of emergency management cycle

Phase	Description	Data Type	Example Data Sets
Pre-emergency (Prevention and Preparedness)	Avoid an incident or intervene to stop an incident from occurring and encompass actions that involve a combination of planning, resources, training, exercising, and organizing to build, sustain, and improve operational capabilities. In this phase governments, organizations, and individuals develop plans to save lives and minimize emergency damage.	User-generated	Twitter (food emergency, earthquake), web traffic (Flu)
		Sensor	Precipitation (PERSIAN, TRMM), evapotranspiration, soil moisture, temperature, vegetation density and water content (MODIS, LANDSAT), groundwater levels (GRACE)
During emergency (Response)	Include immediate actions to save lives, protect property and the environment, meet basic human needs, and preserve business operations.	User-generated	CDR, Flickr, Twitter
		Sensor	Imagery(LANDSAT, MODIS, Geoeye) thermal (LANDSAT, MODIS), radar (RADARSAT-1, CARTOSAT), spatial video
Post-emergency (Mitigation, Recovery)	Design recovery programs to assist victims and their families, restore institutions to suitable economic growth and confidence, rebuild destroyed property, and reconstitute government operations and services affected by emergencies. Recovery activities continue until all systems return to normal or better.	User-generated	CDR, emergency call content, Facebook
		sensor	Night-time Lights (NTL), Imagery, thermal, Radar, spatial video, Temporal Flood Inundation Mapping (GIEMS)
		institutional, public	GCM (Global Climate Model), Transportation data (subway, bikeshare), census, Worldpop, Open Cities

information into decision support requires domain expertise, automated data retrieval, and analytical and computational techniques, and visualization, mapping and decision tools to unveil trends and patters within and between these very large environmental and socio-economic datasets. The significance of big data is growing and expected to close both information and timeliness gaps that limit capabilities to plan, mitigate, or adapt to environmental hazards and change. But various National Meteorological and Hydrological Services and other stakeholders have no means to analyze and utilize effectively the new big data load that is present today and will continue to grow rapidly in the future.

While there is a variety of big data available for each phase of the emergency cycle, understanding issues of scale, granularity, ambiguity, accessibility, representation, and privacy are all key in using big data information correctly and ethically. It is important to understand how to combine data from different resolutions and temporal scales with various emergencies. However, when analysing urban flood risk, high resolution and 3-D imagery is significant to estimate elevation and urban cover to understand where water will flow and pool (Preston et al, 2011, National Academy of Sciences reports People and Pixels 1998, and Tools and methods for estimating populations at risk from natural disasters and complex humanitarian crises, 2007).

Other challenge is difficulty separating the signal from the noise. Selecting the proper algorithm and quantitative metrics to discover precisely robust trends is important, as is understanding that big data analysis a lot demonstrates correlation rather than causation. In the 2010 Haiti earthquake aftermath, social media data production was only weakly correlated with destruction; besides, emergency services faced challenges making SMS information actionable.

It is also well-known that existing big data is not free and public. While Facebook has an open API to access its data, access to Twitter's data stream can be expensive. Gaining access to CDR data requires an agreement with each provider. Some business data is free to view but not download (Zillow, Trulia), and other data can be purchased (Experian Real Estate data, ESRI business analysis). Some satellite data is free (Landsat, MODIS, SRTM), and others for sale (LiDar data, GeoEye, etc.). Access to the required computing power to analyse data can be an issue, but cloud computing and open source software removes some of those barriers.

Furthermore, big data is emerging together with a number of downsides and risks that demand scrutiny. Both academics and practitioners have raised concerns about the representativeness of big data in emergency management. Some big data sources may be representative of particular segments of society, but may not be generalizable to society as a whole (Currion, 2010). For instance, social media data in the wake of Superstorm Sandy were more highly concentrated in less-impacted areas of New York City, rather than in neighbourhoods in south Queens. The platforms on which big data multiply streamline the production and spreading of untruths, too. While means for preventing this are strengthening, practitioners should continue to be cautious about untrustworthy or unconfirmed information, a gradually more challenging task in the big data perspective.

Moreover, privacy and security issues have been a large concern in big data. While data sets that could identify individuals are frequently anonymized (e.g. call record data) even the best attempts to coarsen the data do not preclude individual identification in some cases. So, we should be aware of the sensitivity of anonymized big data sets.

Especially in times of crisis, clear, complete and quick information is needed. 'Disasters are threatening and highly dynamic situations, marked by high levels of information need and low levels of information availability (Shklovski, 2010). Research shows that advances in information and communication technologies enlarge the possibilities for people to seek, get and send information about their situation, feelings and capacities in critical situations. Similarly, the information technology is an essential factor for streamlined search and rescue actions during disasters, but also for sufficient preparation and recovery, stimulated and organised by the authorities.

Shelton et al. (2014) have investigated Twitter activity in the wake of Hurricane Sandy in order to demonstrate the complex relationship between the material world and its digital representations and further described that any analysis of user-generated geographic information must take into account the existence of more complex spatialities than the relatively simple spatial ontology implied by latitude and longitude coordinates.

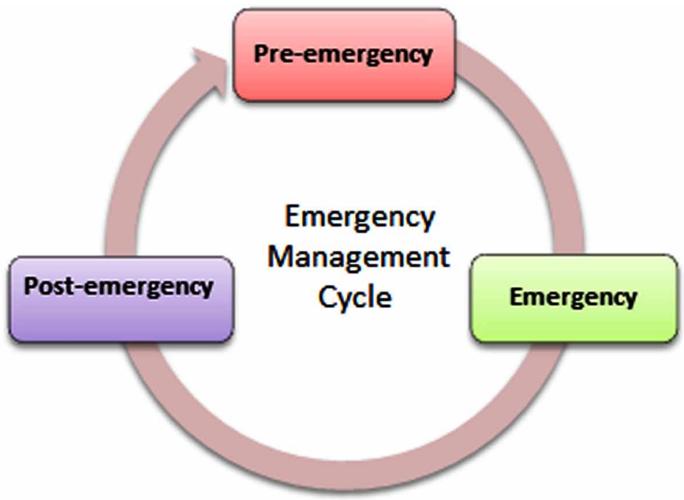
The categorization is based on multiple resources and recognizes that big data can assist before, during and after an emergency, often via a cyclical process, as indicated in Figure 1.

2.1 Pre-Emergency Phase

Big Data analysis can help significantly to the preparation of crisis management. Through the data analysis can be done recognizing the dangers and to provide a sound strategic approach by the respective managers of the crisis. Big Data analysis can also guide the proactive deployment of resources to fully cope with an impeding type of disaster.

Social media data can be used for making diagnoses of vulnerabilities in systems and infrastructures. Besides, some applications can facilitate the warning of citizens in the period before a threatening crisis occurs. For instance, surveys of the American Red Cross show that a large part of the population in the United States is interested in technologies, applications or simply receiving emails for emergency communication. Information about the location of food and water, shelter locations, road closures, the

Figure 1. Emergency management cycle



location of medical services and about how to keep safe during emergencies are high rated by around a half of the questioned people. On the other hand, more than a half of the citizens who is using social media say they should post relevant information on their sites or applications during periods of crisis (Page et al, 2013).

Information derived from the analysis of Big Data can help to anticipate crises or at least reduce the risks that would arise from a disaster the major crisis effect. One example is in a big earthquake harm arises in telecommunication networks leading to interruption of communications, also has been observed a large number of blackouts. There exists a need to study this data for optimization of the civil infrastructure to avoid this crisis effects.

2.2 During Emergency

In emergency situations, big data can be used to provide situational awareness both to authorities and members of the public using information coming from scientists, private organisations and members of the public. This further demonstrates that members of the public are key stakeholders in the big data and emergency management ecosystem.

Big Data analysis in real time can identify which areas need the most urgent attention from the crisis administrators. With the use of the GIS and GPS systems, Big Data analysis can assist the right guidance to the public to avoid or move away from the hazardous situation. Furthermore analysis from prior crisis could help identify the most effective strategy for responding to future disasters.

Moreover disaster-affected communities today are increasingly likely to be 'digital communities' as well – that is, both generators and consumers of digital information. This is obvious from the massive amounts of data being generated by members of the public in emergency situations. However, as well as gathering information from members of the public via social media, this tool can also be used to push information and share information with members of the public to aid in response during an emergency. This is intensely important, as members of the public often act as first responders in emergency situations, well before aid or assistance is available. Many of the systems, including early warning systems, situational awareness systems and training systems also include systems for disseminating information to members of the public, or are particularly intended as collaborative information sharing platforms. Taking emergency mapping as a specific example, the system can be set up in a matter of hours, long before humanitarian or other organizations can arrive. As such, the information can be used to enable members of the public to meet one another's needs in the gap between the incident and the official response.

Real-time big data analysis can substantially enhance various disaster response aspects. First, it can help emergency response personnel to identify areas that need the most urgent attention. This could be areas where there are several people or critical resources. It could also be areas where there may be triggers for other hazards. Second, real-time monitoring and situation analysis can assist emergency response personnel in coordinating their actions to optimally handle a disaster situation. This also includes guidance to the public in taking the best routes to move away from a disaster in order to prevent congestions or causing people to move by mistake to a more hazardous situation. Third, big data analysis from prior incidents can help identify the most effective response methods for various situations and enable the development and deployment of assistive infrastructures for effectively responding to future disasters.

2.3 Post-Emergency Phase

When the recovery activation will gradually start, the infrastructure would provide a big data source. The big data analysis is sharing useful information for recovery procedures about provision of relief supply, volunteer coordination and logistics during the crisis. The mechanisms and approaches for continuous adaptation to the change of demands with the limited resources should be a vital issue for the big data infrastructure.

3. PERTINENT RESEARCH ISSUES

In order to take advantage of big data analysis for efficient emergency management, the core infrastructure must offer exceptional quality of service (QoS). While the QoS requirements may differ for diverse emergency situations, we sketch some instances.

In view of the urgency of the response activities when dealing with emergencies, it is vital for the infrastructure to provide real-time performance. This includes real-time data analysis to precisely predict the impact of an imminent hazard as well as the effective means of responding to the emergency. It additionally involves real-time communication to ensure that accurate data are collected about the situation, such as the location of affected individuals, etc. Real-time communication is also desirable to safeguard that different emergency response teams can coordinate their actions in optimally responding to a disaster (Castillo, 2016).

With the seriousness of emergency response situations, it is crucial to guarantee that the service will be particularly reliable and accessible regardless of the adverse conditions during such circumstances, involving tangible damages, power outages, floods, etc. Consequently, the big data storage, analysis, and communication services must be able to operate regardless of adverse conditions. Computing and sensor resources can be deployed at various geographical locations and communication methods can be used to ensure uninterrupted access to the data.

It is also vital to ensure that the big data supporting infrastructure is sustainable with the evolving nature of emergencies and even emergency response tactics. It is also important to ensure that the service meets high levels of security, such as privacy, confidentiality and assurance that the information used to direct the response to an emergency is proper and not distorted.

Mostly, when we consider the usage of infrastructures, its efficiency, reliability and dependability are key segments. In general, big sensing data are stored in the cloud (Fazio et al, 2015). Yet, in emergency situations, it might not be able to access to the cloud from emergency areas. Thus, it is vital to consider the efficiency and reliability for not only the cloud but also the sensing rims. We also need to consider failures of communication lines in designing and delivering mission-critical services,. Further, it is necessary to ensure that the data acquisition and analysis measures are greatly dependable regardless of the failures of various processing and communication units. Given the distributed nature, it can be difficult to identify which units have failed. Hence, the processing and communication infrastructure may need to be augmented with dependable on-line system health monitoring capabilities to enable the rapid identification of faulty components and the activation of redundant substitute units to ensure correct and apt accomplishment of the big data analysis under emergency situations. Assessment of the big data analysis algorithms is needed to determine the confidence in the correctness of the results of the analysis, including predictions and recommendations for optimal response and recovery actions. Simulation platforms can be used for evaluating and verifying new algorithms and procedures. Different algorithms can be evaluated and compared by applying them to a set of benchmark scenarios and test-beds for which the correct results are known.

3.1. Security and Privacy

There are several significant challenges in information security such as information quality, spam filtering. Crowd-sourcing is among the most effective methods for filtering spam. Nevertheless, due to innate delays crowdsource approach may not be precisely appropriate in emergency situations. In order to enhance the information quality level is to timestamp and location-stamp each message (e.g., enabling the locations feature allows Twitter to show your followers the location you are Tweeting from as part of your Tweet), thus allowing more complete authentication of the data beside its content and correlation with other information sources.

In particular, real-time mining of indirectly self-reported and surveillance information harvested from aggregates of Twitter and other social network feeds can offer useful data and insights about unfolding trends and emerging crowd behaviours at times of crises (Kamel Boulos et al, 2010). However, such (raw) data obtained from Social Web feeds often contain variable amounts of “noise”, misinformation and bias and will commonly require some filtering and verification by both machine-based algorithms and human experts before becoming reliable enough for use in decision-making tasks.

Continuous big data analysis for streaming data, such as output of sensors, results of crowdsourcing, etc., must be enhanced with anomaly detection mechanisms to identify data that may be incorrect due to sensor failures, security attacks, etc. In this aspect the uncertainty quantification method must be constructed with integrated machine learning methods. Machine learning can also help in reducing the chaos in big data storage and analysis by replacing large amounts of data by precise equivalent inference rules. It will also speed up the analysis under emergency situations since the rules can be used to promptly perform the prediction analysis as well as to identify optimal response strategies.

Other issue is about sharing big data and information. Big data producers are mostly reluctant to share information. This problem needs to be tackled for the mission-critical real-time applications such as emergency response, since sensors are collecting such data automatically and appropriate decision making can be achieved through automated integration and sharing, with proper security and privacy assurances. However, in this regard several legal issues need to be solved. For many kinds of emergencies (e.g., flood), there are reliable sources of information. Alas, there are also other kinds of emergencies for which we do not get trustworthy sources from dedicated sensors. Hence research in such direction should emphasis on how people perceive information in social media and how they contribute information in social media.

3.2. Noise & Big Social Data

While traditional event detection approaches assume that all documents are relevant, Twitter data typically contains a vast amount of noise and not every tweet is related to an event[2]. The source of noise in open source information can be divided into intentional and unintentional. Intentional sources of noise and misinformation are generated by cyber-attacks designed for illegitimate financial returns or intentional advantage in a conflict. Social media technologies can facilitate the spread of false information as well as the spread of counter information that attempts to correct the false information, but how to take advantage of these technologies to reduce the spread of misinformation and at the same time increase the spread of useful information such as alerts and warnings is the major concern. The source of unintentional noise is a precipitously varying environment. Linking the information to a correct topic or region of space-time continuum is critical when the environment changes rapidly. Thus improving the quality of information in social media (i.e., filtering out the big noise in big data) is a huge challenge and it involves several prominent issues.

4. SMART TECHNOLOGIES AND BIG DATA

4.1. Crowdsourcing

When a large emergency occurs, it is an incredible challenge to fulfil the information needs of humanitarian responders. Specially, access to latest data on the physical layout of the affected area, the location of critical infrastructure and services is imperative. Likewise, to develop the situational awareness that is desirable to act, responders need information on a kind of assistance required (Stanton et al., 2001). Therefore, maps are of immense significance during crisis response (Meier, 2015).

Crowdsourcing connects unobtrusive and ubiquitous sensing technologies, advanced data management and analytics and novel visualization methods, to create solutions that improve urban environment, human life quality, and city operation systems. Nowadays, no countries, no communities, and no person are immune to urban emergency events. It is important to detect, resistant, and analyse these real time urban emergency events to protect the security of urban residents (Zheng et al, 2016). The crowdsource systems can be useful in mass emergencies to allow people to gather information, report information, volunteer to help, ask for help, or to re-broadcast useful information. Citizens can use this information to determine whether they should follow an evacuation order while government agencies can use this information to determine the allocation of resources or to get an overall sense for the status of a region or city.

For example, during the night of 8th to 9th of June 2014, the storm hit Antwerp after following a destructive path between Gent and Antwerp. Data from fire services was extracted and uploaded to google maps and images from social media were collected. Figure 2 presents a screenshot of flood data from emergency service.

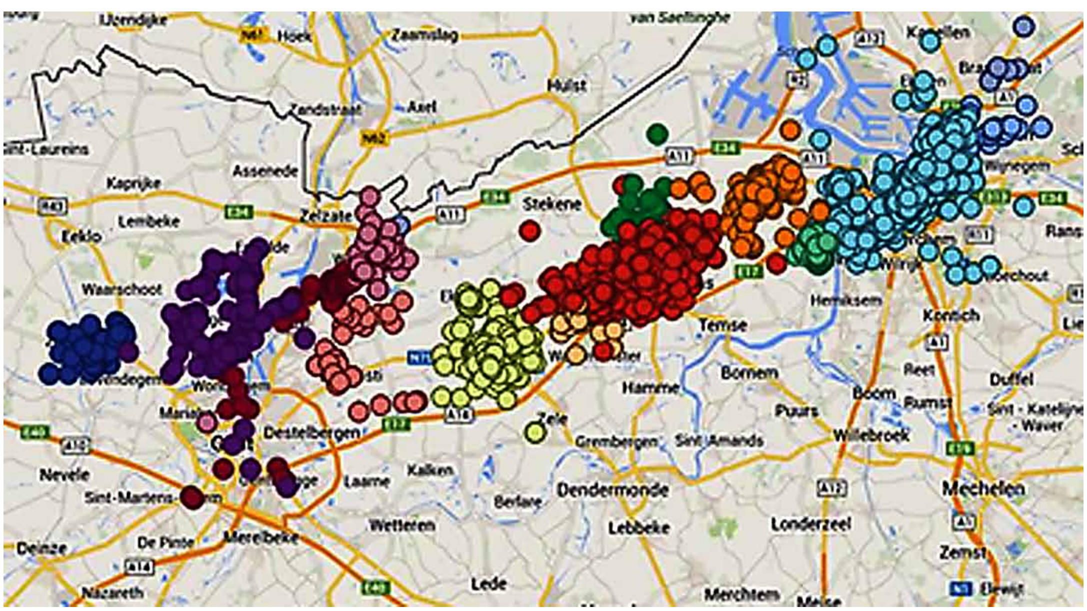
Crisis mapping is the real-time collection, display and analysis of data during a crisis, usually a natural disaster or social/political conflict. Crisis mappers achieve big data analytics and data mapping in order to gather insights about what and where emergency events are occurring on a real-time basis.

Crowdsourcing of information gathering and the sharing of the analysed results by individuals are certainly dominant ways to get the real-time monitoring of the quickly changing situations influenced by lots of unforeseen events. However, gathering information from crowd is challenging. If the collection of data requires users' interaction with the application, then we need to provide users with attractive incentives. Another point is how to deliver the information in a right way in an emergency. Some information may trigger a panic, which may cause further troubles. Nevertheless privacy protection is always a significant problem. However, in order to save as many lives as possible during a disaster, privacy protection policy may be dynamically changed during emergency by emergency services, or even by volunteers. The emergency system should also support the real-time dynamical changes of the policy and the sharing of such information.

4.2. Cyber-Physical-Social Systems (CPSS)

The last decade has seen human factors becoming gradually crucial in computing systems. Therefore, by integrating human factors as part of a system, a cyber-physical-social system (CPSS) encompasses not only cyberspace and physical world, but also human knowledge, mental capacity, and sociocultural

Figure 2. Geographical spread data from fire services



elements. Just as the Internet has transformed the way that people interact with information, CPSS will transform the way people interact with every computing system and create new revolutionary science, technical capabilities for better quality of life. Cyber-Physical-Social systems tightly integrate physical, cyber, and social worlds based on interactions between these worlds in real time. This area is a new research and development field that requires further development of models, methodologies, and theories for efficient interaction between physical, cyber, and social worlds. Cyber-Physical-Social systems rely on communication, computation and control infrastructures commonly consisting of several levels for the three worlds with various resources as sensors, actuators, computational resources, services, humans, etc. (Zeng, 2016; Vardi, 2011; Sheth, 2013). Operation and configuration of CPSS require approaches for managing the variability at design time and the dynamics at runtime caused by a multitude of component types and changing application environments.

Building CPSS for non-emergency conditions is already difficult; yet, building CPSS for emergency management and emergency response is significantly more challenging. Two key important challenges are how to tackle the exponential and multidimensional complexity of their operating environments and how to meet the strict design requirements of CPSS that are necessary for such environments.

Moreover, understanding social theories is an important step toward building systems that interact smoothly with people. These theories are crucial in the design and development of physical-cyber-social systems. For example, a system that interprets various social interactions can be used to capture images/videos. One of the key challenges in dealing with social systems is how to maintain the privacy of participants. Privacy becomes an essential component and it is crucial for wide adoption of physical-cyber-social systems. This is partly due to the fine-grained information collected from sensors and its correlation with behaviour patterns that would reveal personal information which may be misused. For example, a smart-meter installation may result in revealing the occupancy of a house. Some of the key challenges when using social data or in general any data processed by physical-cyber-social computing are listed below.

- Social bots: there are attempts to simulate and flood the social data generated automatically by programs that try to emulate human behaviour. Such a source of information should be used carefully and separated from rest of the social data.
- 'Twitter' data is not always reliable (i.e. requires careful consideration): while social data is available in massive scale (e.g., around 500 million tweets a day), the data is often very noisy, informal, and unevenly distributed.
- Assessing the relevance of Twitter to a problem: not all studies can be done on Twitter data since the nature of data and the social behaviours have a great variance on Twitter.
- Sometimes there is an assumption that data is available at all times: theoretically, there is data available related to various events. However in reality, it may be very hard to find sensors and their observations on Twitter and the Web in general. Choosing the appropriate data source is an important challenge in the context of physical-cyber-social systems.
- Understanding the feedback mechanism: social scientists need to understand the feedback mechanism that exists between the physical world and the social world interactions. This is a challenging and important task to gain insights into systems that involve the social component.
- Data biases are crucial: social scientists should consider data biases carefully with the availability of massive data from social networks such as Twitter.

- Combining reactive vs. non-reactive data: reactive data are those collected by social scientists through surveys and questions. Non-reactive data are those collected by sensors on a continuous basis.

4.3. Service and Cloud Computing

Cloud computing – a long held dream of computing as a utility – is a promising way that shifts data and computational services from individual devices to distributed architectures. Cloud computing provides a convenient tool for crisis response teams to collaborate and share information no matter where the team members are located. Depending on the type of crisis, there may be differing security requirements for the information, and this can impact how the cloud computing is managed or whether additional security measures should be in place. The big data produced during emergency (e.g., by sensors, and social media) have to be collected, integrated, and delivered to big data consumer applications to achieve their new functionality. In addition, emergency management research requires the integration of emergency data with many other big data sources including, but not limited to mapping, land survey, environmental, satellite imagery, population and part disaster datasets; as well as models for climate, geomorphology and hazard spread forecasting.

One of the benefits of cloud computing is that information and operations are hosted in well protected data centers. Leading cloud providers keep information on thousands of systems and in several locations. Redundancy, availability and reliability are hallmarks of cloud computing, so that users can access your information rapidly, no matter where they are located. For example, Amazon and Microsoft have data centers all over the world, with enormous processing power and storage.

Using an analogy with cloud computing and service computing, the big data infrastructure for disaster management is divided into three layers (Pu and Kitsuregawa, 2013). Development of an architecture for each layer of the big data infrastructure for emergency management should consider the distributed nature of the data, the heterogeneity in data sources formats (structured and non-structured), protocols and semantics, the need to meet real-time constraints despite its volume, and quality of data sources.

Disaster recovery is not a problem for cloud service providers but every organization that uses cloud (Chang, 2015). If data are irretrievably lost, this may have negative impacts on the organization affected such as financial loss and loss of time to reproduce or regain data.

(Armbrust et al. 2010) define technical challenges and the security issues to be resolved for cloud computing and also big data. One aspect is ownership and privacy of the data stored in the Cloud. Data in the cloud should have a clear ownership definition, and not be shared or accessed by users without authorization. Legal and data compliance fulfilling governments and regional/international laws need to be met.

(Chang, 2015) demonstrates a multi-purpose approach for disaster recovery, since the existing literature suggests only single approach has been adopted. Chang's approach can meet the requirements for big data systems and services that can fulfil requirements for volume, velocity, variety, veracity and value, with all data restored and updated in four sites. This is particularly useful for service providers who manage cloud data centers.

Clouds are secure and yet adaptation of authenticity, encryption, and meeting security software regulation large concern about secure can be put aside. Besides, the cloud is not in one place; hence the risk of systems failures substantially decreases. In the case of cloud computing, recovery costs are substantially lower since only local computers used to access the Internet are at risk and user data and cloud servers are protected far from the emergency site. In the case of an emergency striking a cloud computing data center, user data will not be lost since suppliers of cloud infrastructure replicate user data and cloud servers across multiple data centers.

5. PROCESSING SOCIAL MEDIA CONTENTS

In above sections we have discussed various issues related to big data processing during emergencies. In this section we will describe our approach to handling social media data. A key challenge when gathering and analysing social media (SM) data is the diversity of different SM services and the presence of different data formats (e.g. a tweet in Twitter or a video in YouTube). Therefore, to allow processing of the heterogeneous SM information we need to standardize the exchange data format.

There are several approaches present that attempt to harmonize multiple SM services on a data level. Social media offers an opportunity to communicate the emergency situation to other citizens or to EMS even though mobile phone or emergency services may be encumbered. During the past few years, various studies were performed focusing on various aspects of social media in emergency management underlining its unceasingly growing importance in this area. For example, to enhance the identification of relevant messages from social media that relies upon the relations between georeferenced social media messages as Volunteered Geographic Information and geographic features of flood phenomena as derived from authoritative data (sensor data, hydrological data and digital elevation models), (de Albuquerque et al, 2015) have proposed an approach to examine the micro-blogging text messages of the Twitter platform (tweets). Several other projects are developing and implementing systems, tools and algorithms performing social media analysis.

This section recaps prominent EU Framework and other project works using social media analysis in the context of emergency management. The following tables summarize the referenced research initiatives based on common aspects of the performed social media analysis. The categorization was done based on respective web sites (e.g., (project) description, screenshots etc.), which does not necessarily include or highlight the full functional range of the systems. Independent of the categorization, all of the mentioned research initiatives are fundamental for situational awareness in emergency management and for monitoring of social media activities.

Now, we illustrate our approach to emergency management IT-system, in the following Figure 3 and describe major modules of the system. The overall objective is a stronger connection between public and emergency services or authorities through social media (Akerkar et al, 2016). In the process, huge amounts of ubiquitous, user-generated content in social media is frequently generated and monitored for emergency related communication. Citizens post messages on social media: amongst those, messages somehow related to emergencies may be present. Once detected, content is gathered and transferred to the analytical phase, where it is pre-processed, analysed according to various data mining analytical approaches. The alerts are generated and then either transferred to routing component for reporting where it can be visualized and interpreted or moves to the communication management module. For instance, The Routing module distributes messages towards responsible emergency service (ES), which use their Command & Control systems or the ES interfaces provided by the system to consume the information. The alert distribution is performed checking continuously for new alerts. Authorities or emergency service personnel can communicate with the public.

Now we will present some key aspects of our IT-system.

Table 2. Social media and emergency management project

Project Name	Project Objectives	Social Media / Data Source	Approaches	Visualization components	Filtering mechanism	Possible Usefulness to our approach
Alert4All [3]	Alert4All aimed at improving the effectiveness of alert and communication towards the population in crises management.	Twitter, blogs	Classification, Support Vector Machine classifier	map	Keyword, tags	Usage of results, especially on how citizens trust information from EMS through different communication channels.
COSMIC[4]	COSMIC project is identifying the most effective ways in which new technologies and applications are being used by citizens and governments.	YouTube, Twitter, Facebook	Classification	list, recommendations and best practices	Topic, information	Mapping the use of current technologies in crises and also mapping the use of emerging applications. Usage of findings on the potential roles and ethics for citizen participation in emergency response.
CrisComScore[5]	The project developed an audit instrument as a tool for ensuring effective crisis communication strategies and implementation.	News media	Text processing	Text messaging	Topic, nformation	The auditing instruments for effectiveness will be one possible measurement method in the analysis phase.
CRISMA[6]	The CRISMA project will develop a simulation-based decision support system, for modelling crisis management, improved action and preparedness.	Data from sensors	Priorisation, Optimisation of response, counter measures and preparedness	GIS based visualization, Real-time environmental data visualization	Information	The auditing of decisions in crisis management will be one possible measurement method in the analysis phase.
ESS[7]	The ESS project developed a common information management and communication platform for supporting the management and coordination of emergency operations.	Real-time sensor data (thermal, video etc.)	Spatial localization, Data fusion	Map, lists	Time, query	Analysis of state-of-the-art technologies for crisis discovery and management and application of existing data fusion methods for developing a data fusion and mediation system.
IDIRA[8]	IDIRA project is focusing on the interoperability of data and emergency procedures in response to large-scale disasters.	Geographic and attribute data, integration of sensor data	Text classification, map, lists	Geo-referenced Visualisation map	Topic, time, information	Methods and technical interoperability standards developed here will influence the integration aspects.
INDECT[9]	The project is developing threat detection tools and generation of data mining and information retrieval applications.	Weblogs, chats, news reports	Relationship mining, machine learning methods for behavioural profiling	Event model	Keyword, query	Consideration of methodologies and algorithms for data & event processing.

continued on following page

Table 2. Continued

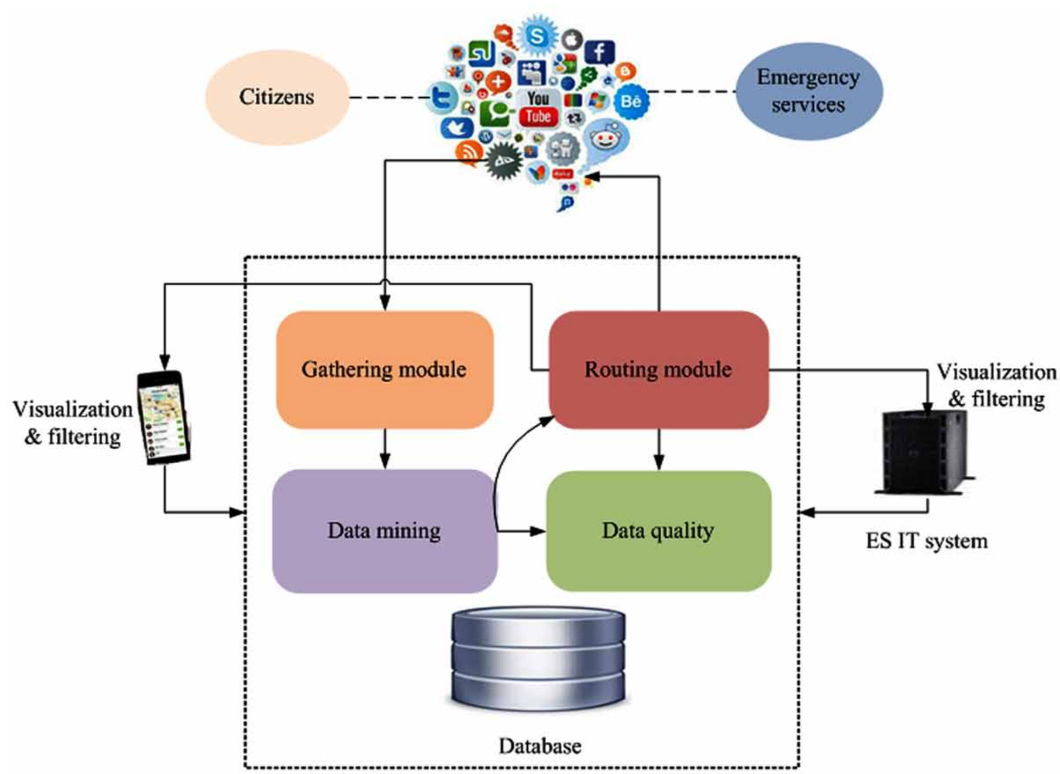
Project Name	Project Objectives	Social Media / Data Source	Approaches	Visualization components	Filtering mechanism	Possible Usefulness to our approach
iSAR+[10]	iSAR+ project delivered the guidelines that enable citizens using new online and mobile technologies to actively participate in response efforts, through the provision, sharing and retrieval of emergency information.	SMS, Twitter, Videos	Multivariete testing, Design of experiments	-	Tags, time, location	Its approach to the dynamics between citizens and EMS in crises, allowing the leverage of EMS' levels of shared awareness and performance, benefiting from citizens' published information.
REACT[11]	REACT has enhanced data by providing associated services that are able to semantically analyse and cluster environmental and crisis management information.	Data from callers and automatic systems	Semantic analysis	GIS based interface	Location, keyword, time, semantic	Usage of the OASIS CAP Protocol to allow interoperability between Emergency Services together with the TSO protocol for a common data ontology.
SocialSensor[12]	SocialSensor is developing a framework for enabling real-time multimedia indexing and search across multiple social media sources.	Facebook, Youtube, Flickr, Tumblr, Google+, Instagram	Clustering of geodata and visual descriptors	Map, timeline, lists	Sorting, zooming	Scalable mining and indexing approaches that taking into account the content and social context of social networks.
WeKnowIt[13]	The system built in this project was based on interviews performed considering emergency management practitioners.	Pictures, Videos, Text	Text processing, Clustering	Map, Timeline	Tag, Time	The results of the layer of social interaction and the massive user feedback layer will be considered as one of the inputs for metrics.
Crisees[14]	Crisees has developed a monitoring tool for social media streams.	Youtube, Twitter	Extension to Sentiment	Map, Lists	Time, Query	Extracting information from social media. Filtering information related to event. Visualising information on maps.
Disaster 2.0[15]	Disaster 2.0 (D2.0) project has explored how EU governments can potentially use Web 2.0 applications and Semantic Technologies in disaster response.	Twitter, Facebook, Ushahidi	Semantic technologies	Map, List	Tag, location, keyword	Use of results in which public utilise web 2.0 and web 3.0 technologies during disasters
Emergency Situation Awareness Platform[16]	The emergency situation awareness platform analysed tweets.	Twitter	Aggregation, Text clasification, Keywords	Map, Tag cloud, Timeline	Slices time, Traffic, Tag	The research results on text classification to identify the impact of the incidents identified.

continued on following page

Table 2. Continued

Project Name	Project Objectives	Social Media / Data Source	Approaches	Visualization components	Filtering mechanism	Possible Usefulness to our approach
SensePlace2[17]	SensePlace2 developed a system for filtering Twitter messages.	Twitter	Named entity recognition	Map, timeline	Time, Tag cloud	Use of different search and filtering facilities to browse through a huge amount of tweets by considering the extracted information.
TEDAS[18]	TEDAS project developed an event detection system for Twitter.	Twitter	Classification & rules, spatial and temporal analysis	Map, Timeline	Location, Keyword, Time	Results to detect and analyze events by exploring rich information from social media
Tweak-the-Tweet[19,20]	Tweak-the-tweet was a crowdsourcing platform. The key difference to other platforms is that this system *'works with the existing social media infrastructure'*.	Twitter	Trust grammar, parsing	Map, Timeline	Predefined hashtags	Parsing algorithm applied to extract information given in combination with hashtags. This extracted information useful to perform keyword-based filtering.
Twitcident[21]	Twitcident project developed a toolbox for filtering and analyzing information from Twitter streams during crisis situations such as fires, storms or other types of incidents.	Twitter	Classification, rules	Map	Keyword, facets	Results related to facet-search interfaces, i.e. on facets extracted in the previous steps, which helps the user to browse through the data to gain a better overview.
TwitterBeat[22]	The TwitterBeat analyzes huge amounts of textual data uncovering the sentiment.	Twitter	Sentiment on topic and location	Map	Zooming	Approaches on sentiment analysis can be used to identify the mood after a disaster, for e.g., for crime prevention.
Twitris[23]	Twitris project presents an opportunity to aggregate social media information. Twitris provides situational awareness by monitoring an event on Twitter at both micro and macro-levels.	Twitter, SMS	Event discriptors	Map, Tag clouds	Spatial temporal theme	Results on context based semantic integration of multiple Web resources and expose semantically enriched social data to the public domain.
Crisees[24]	Crisees has developed a monitoring tool for social media streams.	Youtube, Twitter	Extension to Sentiment	Map, Lists	Time, Query	Extracting information from social media. Filtering information related to event. Visualising information on maps.

Figure 3. EmerGent IT-system



5.1. Data Enrichment

Data enrichment refers to processes used to enhance, refine or otherwise improve raw data. Various studies indicate that extraction of relevant information is a major challenge (Chaudhuri 2012), (Abel 2011). Different circumstances require different assessment methods (Reuter 2015), and different data or meta-data. We can distinguish between *source-based* and *computation-based enriched data*: source-based data is either directly given by the raw data, or not provided, and therefore requires no further computation besides extraction. The actual source-based data varies among different social networks such as Facebook, Google+, Instagram, Twitter or YouTube and is challenged by different technical and business-oriented limitations (Castillo, 2016). Provided data includes date, time, sender, title, tags, keywords, comments, replies, answers, number of views, dislikes, retweets, shares, age, gender, location, education, uploads, watches, total posts and real name. In addition to source-based enriched data, computation-based enriched data, which requires one or multiple steps of algorithmic computation, is vital. Whilst some of the computations can be done on the local server, others may require the invocation of remote APIs. Computationally obtained data includes language detection and sentiment analysis. The provision of enriched data is helpful due to the fact that situation assessment has been shown to be very subjective (Rizza et al, 2013). Thus, information needs depend on personal feelings, experience and the situation itself, wherever information is gathered and analysed, and information systems are implemented to support this task. However, there is a usual challenge on implementing systems to allow both the automatic selection of relevant data and the potentials for end-users to adapt automation and enable tailorable quality assessment according to their requirements. Enriched data supports us to tackle this challenge.

5.3. Semantic Issues

Semantic technologies have the capability to help us cope better with social media data overload. Applying semantic technologies to represent information can provide exceptional means for effectively sharing and using data within different organizations. Using highly structured, self-descriptive pieces of information, interlinked with multiple data resources can help develop a unified and accurate understanding of an evolving scenario. This provides an excellent framework for developing applications and technologies that are highly generic, reproducible and extendible to different regions, conditions, and scenarios. In addition, the semantic descriptions of data can enable new forms of analyses on this data, such as checking for inconsistencies, verifying developments according to planned scenarios, or trying to discover interesting semantically meaningful patterns in data. Such analytics can be performed either in real-time as the scenario unfolds, e.g., through semantic stream processing and event detection techniques, or as an after-action analysis to learn from past events.

On-going research (Galton and Worboys, 2011), (Grolinger et al, 2011) has shown that the need of a common understanding of concepts within and across domains is important to avoid misunderstandings. However, that it is generally accepted to build an ontology from scratch, which does not tap the existing potential of relevant, domain-related knowledge bases. Thus ontologies are often implicitly tailored to a specific need. The ontology can be used to apply semantic analysis on gathered data from SM. This includes the application of further data mining methods to detect patters, incidents or unusual events as well as the detection of correlations. Another advantage of this approach is that emergency services are not essentially required to deal with tweets or posts and may work with domain-related information. To facilitate information exchange with external systems, domains it is necessary that new developments build upon existing standards also on conceptual level. Hence current information models like FOAF, SIOC or MOAC have been considered to construct an ontology that associates information from SM with domain knowledge. Thus one can reuse and extend existing information models in order to combine extracted emergency related content with social media data.

The IT-system is building a large elastic data store in the cloud. It will comprise semantic data stored as RDF for OWL. In order to handle data of enormous size we expect a requirement for parallel computation, subdividing information and execution between different machines that work in the same network. The main challenge to storing RDF objects in NoSQL databases is to find the right way to represent graph inside them. Different studies have tried to use HBase (based on Hadoop) as a NoSQL database coupled with a semantic data framework like Apache Jena (Khadilkar et al, 2012).

We will explore a mixture of NoSQL plus a purely semantic database. For data gathering, a NoSQL solution possibly based on MongoDB is able to fulfil requirements in terms of performance and scalability. We are upholding in the semantic storage only the data that will be analysed (execute queries and apply data mining techniques), and creating a parallel storage for data that is already processed.

5.4. Data Mining

Data mining research has effectively created numerous methods, tools, and algorithms for handling large amounts of data to solve real-world problems. Many standard DM techniques have been developed and successfully implemented across a range of applications (Akerkar and Lingras 2008). However, mining SM comes with a set of unique challenges which have been the focus of much research in recent years (Gimpel et al 2011), (Imran et al 2013), (Castillo 2016). Primary objectives of the data mining process are to efficiently manage large-scale data, extract actionable patterns, and obtain insightful knowledge. Because social media is widely used for various purposes, huge amounts of user-generated data exist and can be made available for data mining. Data mining of social media can expand our capability of understanding new phenomena due to the use of social media and improve business intelligence to provide better services and develop innovative opportunities. In data mining data is transferred into information that needs to be understood in a domain-specific context.

Standard natural language processing (NLP) tools usually fail when faced with the non-standard, untidy language frequently found in SM (Eisenstein 2013), (Liu et al 2012). Additionally, scalability issues are present with SM data. During an emergency the mining tools must be able to capably deal with increased volumes of data. By analysing the enriched data and using NLP techniques on any textual content, combination of mining techniques consolidates multiple SM messages into an information-rich event. However, all together we need a seemly methodology to model any event information.

5.5. Data Quality Issues

Data quality is one of strongest barriers when using citizen-generated content through social media in emergency management. Indeed, issues of reliability, quantification of performance, deception, focus of attention, and effective translation of reported observations/inferences arise when emergency managers start engaging their organisational mechanisms to respond to the disaster. Thus, with the empowerment of the general public and the abundance of information on SM, fostering data quality (DQ) is central for decision makers to achieve an effective and efficient outcome in the emergency response (Jensen 2012). A challenging issue in this domain is to determine how to generate, score, update and represent data and information quality cues to support operators to reason under uncertainties and improve their understanding about an ongoing situation.

Our approach follows the general quality literature by viewing quality as the capability to ‘meet or exceed users’ requirements. Common examples of DQ dimensions are accuracy, completeness, consistency, timeliness, interpretability, and availability. Over the last decade, many studies have confirmed

Table 3. DQ diamensions

Intrinsic dimensions	Contextual Dimensions	Representational Dimensions	Accessibility Dimensions
Believability Accuracy Objectivity Reputation	Value-added Relevance Completeness Timeliness Appropriate amount of data	Interpretability Ease of understanding Representational Consistency Concise representation	Accessibility Access security

that DQ is a multi-dimensional concept and its evaluation should consider different aspects. Despite the multidimensional nature of DQ, it is nevertheless a single phenomenon. DQ dimensions are inherently dependent on each other. For example, to get more accurate information, more time might be required. Accessibility and security are also dependent on each other.

In our research, we are analysing prominent frameworks and develop a DQ assessment technique and reconciliation concept for emergency information. This allows us to assign numerical or categorical values to DQ criteria for information and then subsequently select and prioritise information accordingly to the specific emergency situation. The assessment of the quality of the content in social media adds a significant layer of complexity over traditional DQ assessment frameworks. Challenges arise in timing issues and evaluating the trustworthiness, completeness and accuracy of the quality of content that has been created by users from different backgrounds, in different situations and for different domains.

5.6. Data Visualization for Emergency Decision Support

There are several kinds of visualization techniques[25] for complex – or even SM datasets: simple lists, spatial and temporal representations or charts and graph-based visualizations. The high-level visualization applies very restrictive filters to keep the amount of data as small as possible. The low-level visualization provides a more detailed view on the data.

In our ongoing research project we are tackling aforesaid concerns, with the help of the proposed IT-system the connection between citizens and emergency Services will be enhanced.

6. CONCLUSION

In this paper we have outlined an outlook of processing and analysing big data streams before, during, and after emergencies. Tools that can be used by many emergency services will have significant broad impact in helping citizens as well as many emergency services and government agencies. Big data is a great global opportunity for emergency management. Big data has already demonstrated its usefulness for both dedicated sensor networks (e.g., earthquake detection during the earthquake) and multi-purpose sensor networks (e.g., social media such as Twitter). However, significant research challenges remain, particularly in the areas of Variety of data sources and Veracity of data content. We have also described our efforts in developing emergency management tool that uses social media to support the management of large scale emergencies. It includes the construction of a big online store of data which will be continuously mined to provide emergency information and alerts. Thus, as we keep confronting emergencies, which have become more common during the last few years, emergency service providers can be more efficient in dissemination of warning messages or/and alerts, better manage citizen's sentiment triggered by the disaster, earn trust of citizens, and enhance authorities and citizen cooperation during emergencies.

AKNOWLEDGMENT

The research has received funding from a grant of the European Union (7th Framework Programme No. 608352).

REFERENCES

Abel, F., Gao, Q., Houben, G., & Tao, K. (2011). Semantic enrichment of twitter posts for user profile construction on the social web. Lect. Notes Comput. Sci., 6643, 375–389. doi:10.1007/978-3-642-21064-8_26

Akerkar, R. (2013a). *Big Data Computing*. Chapman and Hall/CRC. doi:10.1201/b16014

Akerkar, R. (2013b). Improving Data Quality on Big and High-Dimensional Data. *Journal of Bioinformatics and Intelligent Control, 2*(1), 155-162.

Akerkar, R., Friberg, T., & Amelunxen, C. (2016). EmerGent Deliverable 3.5: User Requirements (Version 2). Paderborn.

Akerkar, R., & Lingras, P. (2008). *Building an Intelligent Web*. Jones and Burtlett.

Appleby, L. (2013). *Connecting the last mile: The role of communications in the great East Japan earthquake*. London: Internews.

Armbrust, Fox, Griffith, Joseph, Katz, Konwinski, … Zaharia. (2010). A view of cloud computing. *Communications of the ACM*, *53*(4), 50–58.

Becker, H., Naaman, M., & Gravano, L. (2009). Event identification in social media. *Twelfth International Workshop on the Web and Databases*.

Castillo, C. (2016). *Big Crisis Data: Social Media in Disasters and Time-Critical Situations*. Cambridge University Press. doi:10.1017/CBO9781316476840

Chang, V. (2015). Towards a Big Data system disaster recovery in a Private Cloud. *Ad Hoc Networks*, *35*, 65–82. doi:10.1016/j.adhoc.2015.07.012

Chaudhuri, S. (2012). What Next? A Half-Dozen Data Management Research Goals for Big Data and the Cloud. *Proc. 31st Symp. Princ. Database Syst.*, 1–4.

Currion, P. (2010). *"If all You Have is a Hammer" - How Useful is Humanitarian Crowdsourcing?* Retrieved from http://www.crowdsourcing.org/document/if-all-you-have-is-a-hammer---how-useful-is-humanitarian-crowdsourcing/3533

Eisenstein, J. (2013). What to do about bad language on the internet. *Proceedings of NAACL-HLT 2013*, 359–369.

Fazio, M., Celesti, A., Puliafito, A., & Villari, M. (2015). Big Data Storage in the Cloud for Smart Environment Monitoring, Procedia. *Computer Science*, *52*, 500–506.

Galton, A., & Worboys, M. (2011). An ontology of information for emergency management. *Int. Conf. Cris. Response Manag.*, 8, 1– 10.

Gimpel, K., Schneider, N., O'Connor, B., Das, D., Mills, D., Eisenstein, J., … Smith, N. (2011). Part-of-speech tagging for Twitter: Annotation, features, and experiments. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short papers, 2*, 42–47.

Grolinger, K., Brown, K., & Capretz, M. (2011). *From Glossaries to Ontologies: Disaster Management Domain*. Academic Press.

Ilrina, S., Burke, M., Kiesler, S., & Kraut, R. (2010). Technology adoption and use in the aftermath of Hurricane Katrina in New Orleans. *The American Behavioral Scientist*, *53*(8), 1228–1246. doi:10.1177/0002764209356252

Imran, M., Elbassouni, S., Castillo, C., Diaz, F., & Meier, P. (2013). Extracting information nuggets from disaster-related messages in social media. *Proceedings of the 10th international ISCRAM Conference*.

Jensen, G. E. (2012). *Key criteria for information quality in the use of online social media for emergency management in New Zealand*. Victoria University of Wellington.

Kamel Boulos, M. N., Sanfilippo, A. P., Corley, C. D., & Wheeler, S. (2010). Social Web Mining and Exploitation for Serious Applications: Technosocial Predictive Analytics and Related Technologies for Public Health, Environmental and National Security Surveillance. *Computer Methods and Programs in Biomedicine*, *100*(1), 16–23. doi:10.1016/j.cmpb.2010.02.007 PMID:20236725

Khadilkar, V., Kantarcioglu, M., Thuraisingham, B., & Castagna, P. (2012). Jena-HBase: A Distributed, Scalable and Efficient RDF Triple Store. *11th International Semantic Web Conference*.

Liu, X., Zhou, M., Zhou, X., Fu, Z., & Wei, F. (2012). Joint inference of named entity recognition and normalization for tweets. *Proceedings of the ACL.*

Meier, P., & Munro, R. (2010). The unprecedented role of SMS in disaster response: Learning from Haiti. *SAIS Review of International Affairs*, *30*(2), 91–103.

Musaev, De Wang, & Pu. (2014). LITMUS: Landslide Detection by Integrating Multiple Sources. In S.R. Hiltz, M.S. Pfaff, L. Plotnick, & P.C. Shih (Eds.), *Proceedings of the 11th International ISCRAM Conference*, (pp. 677-86). Academic Press.

Page, S., Freberg, K., Saling, K., & Model, E. M. C. V. (2013). A Comparison of Relevant, Timely Message Strategies for Emergency Events. *Journal of Strategic Security*, *6*(2), 20–31. doi:10.5038/1944-0472.6.2.2

Porto de Albuquerque, J., Herfort, B., Brenning, A., & Zipf, A. (2015). A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International Journal of Geographical Information Science*, *29*(4), 667–689. doi:10.1080/13658816.2014.996567

Pu, C., & Kitsuregawa, M. (Eds.). (2013). J*ST/NSF Joint Workshop Report on Big Data and Disaster Management*. Technical Report No. GIT-CERCS-13-09. Georgia Institute of Technology, CERCS.

Reuter, C., Ludwig, T., Ritzkatis, M., & Pipek, V. (2015). Social-QAS: Tailorable Quality Assessment Service for Social Media Content. *Proceedings of the International Symposium on End-User Development (IS-EUD)*. doi:10.1007/978-3-319-18425-8_11

Rizza, C., Pereira, Â., & Curvelo, P. (2013). Do-it-yourself Justice-Considerations of Social Media use in a Crisis Situation: The Case of the 2011 Vancouver Riots. *Proceedings of the Information Systems for Crisis Response and Management (ISCRAM)*, 411–415.

Sheth, A., Anantharam, P., & Henson, C. (2013). Physical-cyber-social computing: An early 21st century approach. *IEEE Intelligent Systems*, *28*(1), 79–82. doi:10.1109/MIS.2013.20

Stanton, N. A., Chambers, P. R., & Piggott, J. (2001). Situational awareness and safety. *Safety Science*, *39*(3), 189–204. doi:10.1016/S0925-7535(01)00010-8

Vardi, M. (2011). Computing for Humans. *Communications of the ACM, 54*(12).

Xu, Z. (2016). *Crowdsourcing based description of urban emergency events using social media big data. IEEE Trans. Cloud Comput.* doi:10.1109/TCC.2016.2517638

Yin, J., Lampert, A., Cameron, M., Robinson, B., & Power, R. (2012). Using social media to enhance emergency situation awareness. *Intell. Syst. IEEE*, *27*(6), 52–59. doi:10.1109/MIS.2012.6

Zeng, J., Yang, L. T., Man Lin, H. N., & Ma, J. (2016). A survey: Cyber-physical-social systems and their system-level design methodology. *Future Generation Computer Systems*. 10.1016/j.future.2016.06.034

KEY TERMS AND DEFINITIONS

Analytics: Using software-based algorithms and statistics to derive meaning from data.

Big Data: Big data refers to the new technologies and applications introduced to handle increasing Volumes of social data while enhancing data utilization capabilities such as Variety, Velocity, Variability, Veracity, and Value.

Emergency Management: The term "emergency management" is used to encompass all of the activities carried out by the federal state and local agencies that are referred to as EMS. These activities have the primary goal of managing hazards, risks, and emergencies of all types.

Data Analytics: The application of software to derive information or meaning from data. The end result might be a report, an indication of status, or an action taken automatically based on the information received.

Scalability: The ability of a system or process to maintain acceptable performance levels as workload or scope increases.

Semi-Structured Data: Data that is not structured by a formal data model, but provides other means of describing the data and hierarchies.

Structured Data: Data that is organized by a predetermined structure.

ENDNOTES

1. Hirschfeld, D. Twitter data accurately tracked Haiti cholera outbreak. Available from http://www.nature.com/news/twitter-data-accurately-tracked-haiti-choleraoutbreak-1.9770.
2. Pear Analytics, "Twitter Study", http://www.pearanalytics.com/wp-content/uploads/2012/12/Twitter-Study-August-2009.pdf,
3. http://www.alert4all.eu/
4. http://www.cosmic-project.eu
5. https://www.jyu.fi/hum/laitokset/viesti/en/research/projects/eucrisiscommunication

[6] http://www.crismaproject.eu/
[7] http://www.ess-project.eu/
[8] http://www.idira.eu/
[9] http://www.indect-project.eu/
[10] http://isar.i112.eu/
[11] http://www.react-ist.net
[12] http://www.socialsensor.eu/
[13] http://www.weknowit.eu/
[14] http://www.dcs.gla.ac.uk/access/crisees
[15] http://repository.disaster20.eu/frontpage
[16] http://www.csiro.au/Outcomes/ICT-and-Services/emergency-situation-awareness.aspx
[17] http://www.geovista.psu.edu/SensePlace2/
[18] https://wiki.cites.illinois.edu/wiki/display/forward/Demo-TEDAS
[19] http://epic.cs.colorado.edu/?page_id=11
[20] http://faculty.washington.edu/kstarbi/TtT_Hurricane_Map_byEvent.html
[21] http://twitcident.com/
[22] http://www.sgi.com/go/twitter/
[23] http://twitris.knoesis.org
[24] http://www.dcs.gla.ac.uk/access/crisees
[25] http://www.interactiondesign.us/courses/2011_AD690/PDFs/Shneiderman_1996.pdf