The use of Volunteered Geographic Information and Crowdsourcing in Disaster Management: a Systematic Literature Review

Flávio E. A. Horita  
Department of Computer Systems, ICMC  
University of São Paulo  
São Carlos/SP, Brazil  
horita@icmc.usp.br

Lívia C. Degrossi  
Department of Computer Systems, ICMC  
University of São Paulo  
São Carlos/SP, Brazil  
degrossi@icmc.usp.br

Luiz F. F. G. Assis  
Department of Computer Systems, ICMC  
University of São Paulo  
São Carlos/SP, Brazil  
luizffga@icmc.usp.br

Alexander Zipf  
GIScience Research Group  
University of Heidelberg  
Heidelberg, Germany  
alexander.zipf@geog.uni-heidelberg.de

João Porto de Albuquerque  
Department of Computer Systems, ICMC  
University of São Paulo  
São Carlos/SP, Brazil  
jporto@icmc.usp.br

ABSTRACT

The number of crisis events around the world has been increasing in the last years and suggests there is a real need to make communities more resilient to them. In addition to providing conventional authoritative data, ordinary citizens and residents in the affected areas are also voluntarily supplying information about the affected areas, in what has been called Crowdsourced or Volunteered Geographic Information (VGI). This paper conducts a Systemic Literature Review aimed at assessing the current state of research in the use of VGI as a source of information to aid the management of disasters. The results suggest there is an increasing body of knowledge of VGI and the way it can improve disaster management. It also reveals gaps in the use of VGI in the research areas of ‘preparedness’ and ‘recovery’, as well as the need for more robust case studies and experimental research to support this promising field.

Keywords
Volunteered geographic information, Crowdsourcing, Disaster management, Systematic literature review, VGI.

INTRODUCTION

In recent years, the numbers of disasters caused by natural hazards has been increasing, and have drawn attention to environmental risks and their potential impact. In particular, there have been approaches targeted at improving community resilience – i.e. the capacity to resist, change or adapt to obtain a new framework for tracing a pattern in the occurrence of disasters (Norris et al., 2008). This has emerged as an important focal point of research studies (De Longueville et al., 2010b; Mediondo, 2010).

The disaster management framework provides an important means of improving resilience and mitigating the impact of natural disasters (Baharin et al., 2009). In managing disasters effectively, it is important for Emergency Agencies to be provided with accurate, timely and complete information, since a slow response based on incorrect data can lead to serious consequences (Ostermann and Spinsanti, 2011; Erskine and Gregg, 2012).

Crowdsourcing or Volunteered Geographic Information (VGI) has emerged in recent years as an important source of information that can support disaster management. These terms are used to reference spatial data that are produced and
disseminated by individuals or non-official institutions, i.e. by ordinary citizens using appropriate tools to gather and disseminate their views and geographical knowledge on the web (Goodchild, 2007; Gill and Bunker, 2012).

Among the advantages associated with VGI, researchers highlight the use of VGI to enhance, update, or complete existing geospatial data sets (Goodchild, 2007; De Longueville et al. 2010b; Goetz and Zipf, 2011). In this manner, VGI applications allow specific information to be collected such as local knowledge, which usually cannot be gathered through traditional data collection processes (Goodchild 2007), and thus empower highly detailed reports to be compiled on local conditions when there are disasters (Zook et al., 2010).

Against this backdrop, the overall research question of this paper is to determine what is the current state of research on the use of VGI in disaster management. This is important to get an overall picture of the available research on the topic, so that lessons can be drawn from the research carried out so far on the subject, e.g. with regard to how VGI can support disaster management in specific types of disaster or the stages in which they occur.

In the pursuit of this goal, this paper conducts a Systematic Literature Review (SLR) to find out the current state of the scientific literature with regard to the use of VGI in disaster management. The remainder of the paper is structured as follows: in Section 2 we define the background concepts used in this study. Section 3 describes the research methodology employed. The results are presented in Section 4 and analyzed in Section 5. Finally, some concluding remarks are made in Section 6.

**BACKGROUND**

**Disaster Management**

A natural disaster occurs when a community is seriously affected by a natural event and the damage is so great that foreign aid is needed (De Longueville et al., 2010b). Some examples are the recent catastrophic events in Haiti, Chile and Pakistan in 2010, Queensland 2010 - 2011, Christchurch and Japan in 2011 which together accounted for 333,944 casualties and economic losses estimated at between 199 and 327 billion US dollars (Erskine and Gregg, 2012).

One way to mitigate the effects of this damage is through the adoption of reactive measures to prevent these events from turning into a catastrophe (Baharin et al., 2009; Poser and Dransch, 2010). Thus, an approach must be adopted to improve the resilience of communities and enable them to resist, adapt or change when a disaster occurs (Mediondo, 2010; Norris et al., 2008).

Disaster management is an alternative to improving resilience and thus avoids or reduces the impact of natural disasters (Baharin et al., 2009). It can be defined as an ongoing process involving a series of activities before, during and after an event, which are separated into four main phases: mitigation, preparedness, response and recovery. Figure 1 shows these phases and their associated activities.

![Figure 1. Phases of disaster management and its activities (Poser and Dransch, 2010).](image-url)

As shown in Figure 1, disaster management involves a cycle of activities that begins with estimating vulnerability to negative disasters and the mitigation of their impacts; it involves preparation at an operational level and, after the outbreak of the event, enables reactive measures to be taken to deal with problems and help the community to recover.
In all these phases, information used by Emergency Agencies (EAs) plays a key role by ensuring the mitigation of the effects. As a result, it is essential that there is accurate, timely and complete information about the current state of the environmental variables, as well as about the scientific predictions of upcoming changes and their expected effects.

Volunteered Geographic Information (VGI)

For years, the growth of spatial data was limited to activities carried out by specialist organizations (Goodchild and Glennon, 2010). Following the emergence of Web 2.0 and the improvement of mobile devices to provide data related to their location, people began to be more involved in this area, not just by examining it, but also by providing data and information, that was in many cases, more detailed and of a higher quality than that provided by official institutions (Goodchild, 2007; Elwood, 2008; De Longueville et al., 2010b; Gill and Bunker, 2012).

This kind of information was defined by Goodchild (2007) as Volunteered Geographic Information (VGI) which is a collection of digital spatial data produced by individuals and informal institutions, i.e. by ordinary citizens using appropriate tools to gather and disseminate their views and geographical knowledge on the web. According to Coleman et al. (2009), this ‘volunteerism’ is an effective way to expand, qualify, and contribute information for humanitarian purposes. This improved the amount of information available about events and the experiences of the community members, using tools like the grassroots maps of the Humanitarian OpenStreetMap Team (HOT) and Wikimapia¹, as well as the crisis maps of Ushahidi (Okolloh, 2009; Ziemke, 2012). Together, the people involved form a new group of players in disaster management who are known as Volunteer and Technical Communities (VTC) or digital humanitarians.

These players have produced a considerable amount of information on the Internet. A study in the U.S. shows that 35% of these users create content and put it online, and 26-34% of them share material published by others (Flanagin and Metzger, 2008). This huge volume of data is a potentially valuable resource. However, research has also pointed out that this has given rise to a number of difficulties and obstacles.

The main issues that are raised concerning the use of VGI are credibility, reliability and quality of VGI (Flanagin and Metzger, 2008; Elwood, 2008). VGI can be regarded as lacking credibility and reliability because it is produced by non-experts in a context that differs significantly from “structured institution-initiated and expert-driven contexts” (Elwood, 2008). In addition, the sheer abundance of this information has also highlighted a problem about its structuring, storage and dissemination, i.e. how a Spatial Data Infrastructure (SDI) can be defined that can address these requirements (Bishr and Janowicz, 2010).

RESEARCH METHOD

A Systematic Literature Review (SLR) is a means of evaluating and interpreting all the studies available in the literature about research questions, area, or phenomenon of interest (Kitchenham and Charters, 2007). Conducting systematic research follows a sequence of well-defined and rigorous methodological steps, in accordance with a previously prepared protocol (Biolchini et al., 2005). According to Kitchenham and Charters (2007), a SLR comprises three steps: planning, conducting and results. In the planning phase, the research questions are defined together with the SLR protocol. In the conducting phase, the planning is executed in the electronic databases and a selection made of the returned studies. Lastly, the returned studies are evaluated and the data extracted. In the next section, there will be a detailed account of the approach adopted for this study based on the steps proposed by Kitchenham and Charters (2007).

Research Questions

The formulation of research questions plays an important role during the planning stage because they will be used as guidelines to define the activities and processes employed in the execution of SLR (Dyba et al., 2007). As the main objective of this study is to identify how information from volunteers is of value in achieving resilience in disaster management, following research questions have been selected:

- **RQ1**: In which phases of disaster management has volunteered information been used?
- **RQ2**: In what types of disasters is volunteered information used?
- **RQ3**: What types of methodologies are employed in research that use volunteered information for disaster management?

¹ [http://wikimapia.org/](http://wikimapia.org/)
The four phases of disaster management defined by (Poser and Dransch, 2010) were used as a benchmark to determine at what stage RQ1 is included. In the case of RQ2, we used the types outlined by Jha (2010) to identify the category, type and subtype of disaster employed in this paper. To answer RQ3, we used the information provided by the reviewed papers themselves about the research methodology employed, e.g. the literature review, case studies, surveys, experiments, and action research.

Search Process

The search process was performed in electronic databases (Table 1 shows these databases) with the aim of identifying papers that are important for this study (Step 1). A search string was defined subjectively (Zhang et al., 2011) to be applied to each of these bases, composed of the terms “Volunteered Geographic Information” and “Disaster Management”. Moreover, at this stage, the string preparation has been carried out to achieve the greatest volume of relevant papers, syntax and coding adopted by each search engines.

Initially, we assessed the title and abstract and applied pre-defined inclusion and exclusion criteria to each paper, and thus obtained a set of relevant papers (Step 2). Following this, the full text of each of these papers was obtained and its contents were critically evaluated. In doing so, the 26 studies were read in full and the inclusion and exclusion criteria applied, which in the end comprised a selection of 21 studies (Step 3). For this step, a form was used to extract the data and integrate the researcher’s answers in the best way possible. Lastly, the main data were summarized and the relevant studies categorized.

RESULTS

320 papers were found in Step 1. Among them, the SpringerLink had the highest return, 183 papers, (57% in total). In Step 2, the papers were filtered and a list with 26 relevant papers was defined. Table 1 displays these data in more detail.

<table>
<thead>
<tr>
<th>Source</th>
<th>URL</th>
<th>Results</th>
<th>Relevant Papers</th>
<th>% to total</th>
<th>% used papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEExplore</td>
<td><a href="http://www.ieeexplore.ieee.org">http://www.ieeexplore.ieee.org</a></td>
<td>8</td>
<td>1</td>
<td>2%</td>
<td>12%</td>
</tr>
<tr>
<td>ACM Digital Library</td>
<td><a href="http://www.portal.acm.org">http://www.portal.acm.org</a></td>
<td>3</td>
<td>1</td>
<td>2%</td>
<td>33%</td>
</tr>
<tr>
<td>SCOPUS</td>
<td><a href="http://www.scopus.com">http://www.scopus.com</a></td>
<td>27</td>
<td>7</td>
<td>8%</td>
<td>18%</td>
</tr>
<tr>
<td>Web of Knowledge</td>
<td><a href="http://www.webofknowledge.com">http://www.webofknowledge.com</a></td>
<td>8</td>
<td>2</td>
<td>3%</td>
<td>25%</td>
</tr>
<tr>
<td>AIS Eletronic Library</td>
<td><a href="http://aisel.aisnet.org/">http://aisel.aisnet.org/</a></td>
<td>36</td>
<td>3</td>
<td>11%</td>
<td>7%</td>
</tr>
<tr>
<td>Science Direct</td>
<td><a href="http://www.sciencedirect.com">http://www.sciencedirect.com</a></td>
<td>55</td>
<td>4</td>
<td>17%</td>
<td>2%</td>
</tr>
<tr>
<td>SpringerLink</td>
<td><a href="http://www.springerlink.com">http://www.springerlink.com</a></td>
<td>183</td>
<td>8</td>
<td>57%</td>
<td>3%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>320</td>
<td>26</td>
<td>100%</td>
<td>8%</td>
</tr>
</tbody>
</table>

Table 1. Electronic database used.

Table 1 provides evidence of the widespread use of Web of Knowledge and ACM Digital Library showing an index of 25% (2 out of 8) and 33% (1 out of 3). During Step 3, paper developed by Ahmed (2011), although only a work-in-progress, provided support for the application of social media for disaster management, and was thus maintained among the relevant papers. Moreover, at this stage, some papers were excluded (Curtis and Mills, 2012; Boulos et al., 2011; Kavanaugh et al., 2012) since they do not use or apply VGI for disaster management; Ghosh et al. (2011) is related to the crowdsourcing carried out by know-how specialists, and Chu et al. (2011) is related to models and strategies for collecting information available on the Internet. Table 2 shows the relevant papers in detail.
<table>
<thead>
<tr>
<th>ID</th>
<th>Author</th>
<th>Research Method</th>
<th>Disaster Phase</th>
<th>Disaster Type</th>
<th>Sharing Media</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>(Pearce, 2005)</td>
<td>Case Study</td>
<td>Mitigation</td>
<td>Not Applied</td>
<td>Workshops</td>
</tr>
<tr>
<td>P2</td>
<td>(Chen et al., 2006)</td>
<td>Case Study</td>
<td>All of them</td>
<td>Not Applied</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>P3</td>
<td>(Schafer et al., 2007)</td>
<td>Case Study</td>
<td>All of them</td>
<td>Not Applied</td>
<td>Geocollaboration Software</td>
</tr>
<tr>
<td>P4</td>
<td>(Ikeda et al., 2008)</td>
<td>Design Science</td>
<td>Mitigation Preparedness</td>
<td>Floods</td>
<td>Lectures</td>
</tr>
<tr>
<td>P5</td>
<td>(Huang et al., 2010)</td>
<td>Case Study</td>
<td>Response</td>
<td>Storm</td>
<td>Mobile Devices, Web Services</td>
</tr>
<tr>
<td>P6</td>
<td>(Sinnappan et al., 2010)</td>
<td>Case Study</td>
<td>Response</td>
<td>Fire</td>
<td>Twitter</td>
</tr>
<tr>
<td>P7</td>
<td>(Poser and Dransch, 2010)</td>
<td>Case Study</td>
<td>Response</td>
<td>Floods</td>
<td>Phone Surveys</td>
</tr>
<tr>
<td>P8</td>
<td>(Goodchild and Glennon, 2010)</td>
<td>Case Study</td>
<td>Response</td>
<td>Fire</td>
<td>Twitter</td>
</tr>
<tr>
<td>P9</td>
<td>(De Longueville et al., 2010a)</td>
<td>Design Science</td>
<td>Response</td>
<td>Forest Fire</td>
<td>WebServices</td>
</tr>
<tr>
<td>P10</td>
<td>(Schade et al., 2011)</td>
<td>Case Study</td>
<td>Response</td>
<td>Floods</td>
<td>Flickr</td>
</tr>
<tr>
<td>P11</td>
<td>(Savelyev et al., 2011)</td>
<td>Design Science</td>
<td>Not Identified</td>
<td>Not Identified</td>
<td>Mobile Devices and Web Services</td>
</tr>
<tr>
<td>P12</td>
<td>(Roche et al., 2011)</td>
<td>Literature Review</td>
<td>Response</td>
<td>Not Applied</td>
<td>Not Identified</td>
</tr>
<tr>
<td>P13</td>
<td>(Yates and Paquette, 2011)</td>
<td>Case Study</td>
<td>Response</td>
<td>Landslide</td>
<td>SharePoint</td>
</tr>
<tr>
<td>P14</td>
<td>(Ahmed, 2011)</td>
<td>Literature Review</td>
<td>All of them</td>
<td>Earthquake, Volcano, Landslide and Fire</td>
<td>Twitter, Facebook, Youtube, Flickr</td>
</tr>
<tr>
<td>P15</td>
<td>(Niko et al., 2011)</td>
<td>Design Science</td>
<td>Mitigation Preparedness</td>
<td>Not Applied</td>
<td>Not Identified</td>
</tr>
<tr>
<td>P16</td>
<td>(Pohl et al., 2012)</td>
<td>Design Science</td>
<td>Preparedness</td>
<td>Not Applied</td>
<td>Flickr, Youtube</td>
</tr>
<tr>
<td>P17</td>
<td>(Vivacqua and Borges, 2012)</td>
<td>Design Science</td>
<td>Response</td>
<td>Fire</td>
<td>Mobile Devices and Web Maps</td>
</tr>
<tr>
<td>P18</td>
<td>(Weaver et al., 2012)</td>
<td>Design Science</td>
<td>Response</td>
<td>Several</td>
<td>Mobile Devices and Web Site Developed</td>
</tr>
<tr>
<td>P19</td>
<td>(Adam et al., 2012)</td>
<td>Case Study</td>
<td>Not Identified</td>
<td>Not Applied</td>
<td>Not Identified</td>
</tr>
<tr>
<td>P20</td>
<td>(Kaewkitipong et al., 2012)</td>
<td>Case Study</td>
<td>All of them</td>
<td>Floods</td>
<td>Twitter, Facebook, Youtube</td>
</tr>
<tr>
<td>P21</td>
<td>(Munro, 2012)</td>
<td>Case Study</td>
<td>Response</td>
<td>Earthquake</td>
<td>Text Messages</td>
</tr>
</tbody>
</table>

Table 2. Summary of the reviewed papers.

**ANALYSIS**

In this section, we will summarize the data outlined in the previous section and show some indicators that were identified and defined to address the research questions.
RQ1: In which of the disaster management phases has voluntary information been used?

With regard to the use of VGI in the different phases of disaster management, we identified a strong predominance of works covering the response phase (11 out of the 21 selected papers). Three other papers discussed the mitigation phase and three were also found in the preparation phase. Figure 2 shows these papers distributed per year.

Moreover, in general, the studies of Ahmed (2011), Schafer et al. (2007), Chen et al. (2006) and Kaewkitipong et al. (2012) were carried out in more than one stage of the disaster management. Savelyev et al. (2011) and Adam et al. (2012) do not identify which phase of disaster management their research was concerned with.

RQ2: In what types of disasters was the voluntary information used?

Figure 3 displays the disasters types where volunteered geographic information is being used with more frequency. As can be seen, the occurrence of floods, earthquakes, fires and storms are very common.

As can be seen in Figure 3, almost 50% of the research in the field entails carrying out activities for managing disasters related to fire and floods, and another 30% were related to storms and earthquakes. These facts can be explained by the growing number of these types of disasters around the world. Some research has been conducted in countries like Taiwan, Japan, Thailand, and South Korea that are seriously affected by tectonic instability which causes these types of disasters. Moreover, on the basis of data collected during the SLR, it was revealed that there is still a gap in the research on data involving the use of volunteers for disasters in Latin America, which have recently experienced severe floods, fires and storms.
RQ3: What are the types of methodologies used for research that aims to use voluntary information for disaster management?

Table 3 shows the quantification of each of the key research methodologies employed by the different papers. As disaster management is an area with a strong field for applications, the majority of the studies analyzed used methods for conducting case studies (12 papers). The design of new processes, techniques and methods following the design science approach was also used by a significant number of works (7 papers). Only a few papers have conducted a literature review to explore papers in the area so far (2 papers).

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Papers</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Study</td>
<td>12</td>
<td>57,14%</td>
</tr>
<tr>
<td>Literature Review</td>
<td>2</td>
<td>9,52%</td>
</tr>
<tr>
<td>Design Science</td>
<td>7</td>
<td>33,34%</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100,00%</td>
</tr>
</tbody>
</table>

Table 3. Methodology of the reviewed papers.

FINAL REMARKS

This paper has conducted a Systematic Literature Review aimed at assessing the current state of research into the use of VGI as a source of information for disaster management. For this purpose, we followed the guidelines proposed by Kitchenham and Charters (2007) and selected papers from seven important literature databases (IEEEExplore, ACM Digital Library, SCOPUS, Web of Knowledge, AIS Electronic Library, Science Direct and SpringerLink).

The findings of this study established that the scientific literature about the use of VGI in disaster management is growing, with a significant growth in the number of publications over the last three years (2010-2012). Moreover, it was found that the predominant research area was disaster response, whilst fewer studies were devoted to mitigation and preparedness, and no study was found dealing with recovery. This can be explained by the fact that response is the most visible part of disaster management and is also more likely to attract the attention of volunteers. However, there is a challenge for researchers which is how to advance knowledge about methods that include VGI in mitigation and preparedness activities such as risk analysis and early warning systems, as well as in the recovery phase of communities, by helping them reorganize their routine, and create mechanisms to prevent disasters from happening again in the future.

This research also showed that VGI is commonly used to manage floods and fires. The prevailing media for sharing VGI was found to be social media (i.e. Twitter, Facebook, Youtube) and mobile devices. Interestingly, very few of the reviewed papers address VGI platforms like Ushahidi², Elva¹, OpenStreetMap, and Wikimapia; this offers an interesting avenue to explore in future research. Furthermore, apart from the fact that most current studies are based on case studies, there is also a broad field for conducting research that employs other methodologies. Action research seems particularly suited to this area, since it seeks to solve current practical problems (e.g. establishing resilience against disasters with the aid of volunteered information) while expanding scientific knowledge (e.g. learning how we can effectively use volunteer information to tackle disasters) (Baskerville and Myers, 2004).

Despite their limitations, the results of this study should be broadly welcome. However, it should be stressed that the final number of selected papers (21) is still quite small. This can be attributed to the fact that we did not include synonyms in the research strings used, or include manual searches in particular conference proceedings, or supplement the search with any other sources apart from the results achieved from the search engines. Future studies could overcome these limitations, e.g. by applying the snowball method (Webster and Watson, 2002) in finding out more valuable papers from the citations of the papers in the initial set, as well as employing other techniques such as the use of this set as a quasi-gold-standard (Zhang et

---

¹ http://www.ushahidi.com/
² http://www.myelva.com/
al., 2011). Finally, it might be useful to include new keywords (e.g. synonyms) to the search string and executing manual searches in selected outlets that are more specialized.

ACKNOWLEDGMENT

The authors would like to express their gratitude to FAPESP (processes 2011/23274-3 and 2012/18675-1) and the Alexander von Humboldt Foundation for their financial support.

REFERENCES


