2019

Crowdsourced social media data for disaster management: Lessons from the PetaJakarta.org project

Robert Ighodaro Ogie
University of Wollongong, rogie@uow.edu.au

Rodney J. Clarke
University of Wollongong, rclarke@uow.edu.au

Hugh I. Forehead
University of Wollongong, hughf@uow.edu.au

Pascal Perez
University of Wollongong, pascal@uow.edu.au

Publication Details

Ogie, R. I., Clarke, R. J., Forehead, H. & Perez, P. (2019). Crowdsourced social media data for disaster management: Lessons from the PetaJakarta.org project. Computers, Environment and Urban Systems, 73 108-117.

Crowdsourced social media data for disaster management: Lessons from the PetaJakarta.org project

Abstract

The application of crowdsourced social media data in flood mapping and other disaster management initiatives is a burgeoning field of research, but not one that is without challenges. In identifying these challenges and in making appropriate recommendations for future direction, it is vital that we learn from the past by taking a constructively critical appraisal of highly-praised projects in this field, which through real-world implementations have pioneered the use of crowdsourced geospatial data in modern disaster management. These real-world applications represent natural experiments, each with myriads of lessons that cannot be easily gained from computer-confined simulations. This paper reports on lessons learnt from a 3-year implementation of a highly-praised project- the PetaJakarta.org project. The lessons presented derive from the key success factors and the challenges associated with the PetaJakarta.org project. To contribute in addressing some of the identified challenges, desirable characteristics of future social media-based disaster mapping systems are discussed. It is envisaged that the lessons and insights shared in this study will prove invaluable within the broader context of designing socio-technical systems for crowdsourcing and harnessing disaster-related information.

Disciplines

Engineering | Physical Sciences and Mathematics

Publication Details

Ogie, R. I., Clarke, R. J., Forehead, H. & Perez, P. (2019). Crowdsourced social media data for disaster management: Lessons from the PetaJakarta.org project. Computers, Environment and Urban Systems, 73 108-117.

Originally published as:
Ogie, R, Clarke, R, Forehead, H, Perez, P 2019, 'Crowdsourced social media data for disaster management: Lessons from the PetaJakarta.org project', *Computers, Environment and Urban Systems*, vol. 73, pp. 108-117.

Abstract

The application of crowdsourced social media data in flood mapping and other disaster management initiatives is a burgeoning field of research, but not one that is without challenges. In identifying these challenges and in making appropriate recommendations for future direction, it is vital that we learn from the past by taking a constructively critical appraisal of highly-praised projects in this field, which through real-world implementations have pioneered the use of crowdsourced geospatial data in modern disaster management. These real-world applications represent natural experiments, each with myriads of lessons that cannot be easily gained from computer-confined simulations. This paper reports on lessons learnt from a 3-year implementation of a highly-praised project- the PetaJakarta.org project. The lessons presented derive from the key success factors and the challenges associated with the PetaJakarta.org project. To contribute in addressing some of the identified challenges, desirable characteristics of future social media-based disaster mapping systems are discussed. It is envisaged that the lessons and insights shared in this study will prove invaluable within the broader context of designing socio-technical systems for crowdsourcing and harnessing disaster-related information.

Keywords: social media; situational awareness; flood mapping; disaster; civic participation; VGI

1. Introduction

Natural disasters are increasing globally, both in frequency and intensity (Crooks & Wise, 2013). The global yearly loss associated with these disasters run into billions of dollars, with thousands of deaths and injuries recorded (Lu et al., 2016). This situation calls for more effective management of natural disasters, particularly in highly vulnerable regions of the world such as the mega-cities of developing nations, where population exposure is high and emergency response is largely constrained by shortage of resources and lack of situational information (Ogie et al., 2017b). Interestingly, social media has emerged as a potentially useful technology to assist in disaster management activities, including speedy detection of socially disruptive events, facilitation of crisis communication, and attainment of situational awareness (De Albuquerque et al., 2015). The phenomenal growth in the richness and diversity of time-critical information shared on social media platforms during natural disasters provides a unique opportunity to harvest large-scale spatio-temporal data of immense value to emergency managers (Granell & Ostermann, 2016).

However, the process involved in harnessing social media for disaster management is quite challenging and there is no single, 'silver bullet' solution in this relatively new and emerging field (Croitoru et al., 2013). While disaster management and social media research is burgeoning, only a few solutions, including Ushahidi, CrisisTracker, Cognicity, TweetTracker, Twitcident etc, have actually been deployed in real-world disaster scenarios (Abel et al., 2012; Holderness &Turpin, 2015; Kumar et al., 2011; Rogstadius et al., 2013). These real-world applications represent natural experiments, each with a unique set of lessons that can potentially help to improve how we apply social media to manage future disasters (Yates & Paquette, 2011). The importance of these lessons cannot be overemphasised, as they provide understanding of the success or failure of systems beyond their technical capabilities, taking cognizance of the complex social world in which they are embedded (Coiera, 2007)

In the present study, the authors share lessons learnt from the PetaJakarta.org (Map Jakarta) project. The PetaJakarta.org project is a social media-based, citizen-driven data collection and disaster mapping initiative that aims to advance the capacity of cities to understand and promote resilience to extreme weather events (Holderness & Turpin, 2015). Distinct from previous publications (Holderness & Turpin, 2015; Ogie et al., 2018b; Perez et al., 2015) that merely convey descriptive information about the PetaJakarta.org project, the present study contributes by providing a constructively critical evaluation of the project, including a myriad of lessons that can potentially help to improve future applications of crowdsourced social media data in disaster mapping. Lessons discussed will reflect upon the social and technical aspects of a 3-year implementation of the PetaJakarta.org system in Jakarta, Indonesia. Understandably, systems designed to harness the power of social media during disaster events function as socio-technical systems (Perez et al., 2015), and as such should be evaluated from that perspective. Both the key success factors (KSF) and the challenges of the PetaJakarta.org system will be explained to help improve the success of future applications. The study will further contribute by describing desirable characteristics of future social media-based flood mapping systems, which when adopted can help to address some of the identified challenges. Some techniques for achieving these system characteristics will be covered. The following section presents a brief review of relevant literature.

1.1. Literature Review

With increasing interest in leveraging the “wisdom of the crowd” during crisis events, a number of studies have previously attempted to share lessons from crowdsourcing experience that highlight the benefits and challenges associated with harnessing crowdsourced social media data for disaster management. The benefits of leveraging crowdsourcing for disaster management have been identified consistently in the literature and they include near real-time data collection, cost saving, enhanced crisis communication, locational cues to where help is needed, collective intelligence and improved situational awareness and ground truth (Gao et al., 2011; Roche et al., 2013; Zook et al., 2010).

However, there are several challenges that militate against social media applications in disaster management. The most crucial of these challenges are those related to data quality, validation and accuracy (Goodchild, 2007). Other major drawbacks include erroneous inferencing of crisis locations from geotagged data, fraudulent or non-genuine reports, tensions between utilising expert knowledge and crowdsourced amateur opinions, false alarms inherent in asserted information, poor processing of duplicate reports (e.g., retweets), scalability and interoperability issues, inequality in access to participatory tools (e.g., technology, skills, education), need to improve platform reliance and stability through a multi-platform approach, inadequate report verification and trust management mechanism, privacy concerns, inadequate features to protect the safety and security of data contributors and relief workers, poor filtering of voluminous data as required for report summarisation, and the lack of predictive capabilities to forecast the demand for disaster relief resources at various locations and time (Gao et al., 2011; Huang et al., 2010; Shanley et al., 2013; Brandusescu & Sieber, 2017; Zook et al., 2010).

Furthermore, Gao et al. (2011) reported that existing crowdsourcing applications, designed to harness social media data, are not robust enough to facilitate collaboration and coordination within and between disparate relief organisations. Based on lessons learnt from using Twitter in Japan's tsunami disaster, Acar & Muraki (2011) highlighted a lack of effective ways to trace and expunge rapidly spreading information that have already been actioned on by responsible authority. They also noted inconsistency and confusion resulting from using multiple synonymous hash tags to share similar information on Twitter. Shanley et al. (2013) and Roche et al. (2013) emphasised on legal and ethical concerns, including intellectual property, data access and distribution, licensing issues and liability arising from negligence in utilising crowdsourced social media data. Another major concern is the need to design for adaptability and complementarity in case there is a failure, for example, in the network infrastructure (Huang et al., 2010; Roche et al., 2013). In addition, Palen & Anderson (2016) discussed the need to consider other issues such as shortage of geotagged data, heterogeneity in data format, inadequacies in population representation due to poor sampling decisions, and shortfall in deriving complete meaning and value from social media data as a result of failure to capture and process associated contextual information.

2. The PetaJakarta.org system

The PetaJakarta.org system is a social media-based urban data collection initiative developed by the SMART Infrastructure Facility of the University of Wollongong, Australia in collaboration with the DKI Jakarta Regional Disaster Management Agency (also known as BPBD DKI Jakarta) and Twitter Inc. (Holderness & Turpin, 2015). The PetaJakarta.org system aims to solicit, gather, sort, and map citizens' reports of real-time flood conditions through social media. The system relies on citizens engaging with social media to report the locations and conditions of flood events in their localities. Citizens' flood reports are then aggregated and used to generate a publicly accessible city-scale map of real-time flood conditions. The flood map is generated by the system using geolocated reports that have been confirmed by the sender to be related to an ongoing flooding event. For more details about the method used in generating the flood map, please refer to the concept of "Big Crowdsourcing" already covered in a previous study by Holderness & Turpin (2015). During disaster situations, such collectively generated information about flood locations and severity enables citizens to navigate safely through cities. Government agencies, especially emergency services, can also be better informed to make optimal and priority-based decisions relating to dispatch of relief materials and rescue operations. Unlike many planning and decision support systems that end as mere technical successes with little or no adoption in actual emergency planning or operations, the PetaJakarta.org system was deployed in a pilot implementation in the city of Jakarta, Indonesia.

Jakarta was selected for the study because it is typical of coastal mega-cities of developing nations, which are desperately in need of innovative solutions to the problem of flood disasters (Ogie et al., 2016). Every year, the city experiences flooding caused by a combination of factors, including changing climatic conditions, physical geography, land subsidence, sea-level rise, shortage of funding, aging and deteriorating flood control infrastructure, poor maintenance of drainage network, unplanned urban growth, slum dwellings and the illegal disposal of solid waste in waterways (Ogie et al., 2018a). The situation is further exacerbated by the lack of reliable and timely flood data needed to enable emergency workers and managing authorities gain situational awareness and respond swiftly to unfolding flooding events in various parts of the city (Ogie et al., 2017a). Interestingly, Jakarta residents are the world's most active Twitter users, sharing disaster situational information on social media to help their families, friends and local communities stay safe (Eilander et al., 2016). The government of Jakarta also promotes social media as a means of communicating with citizens during disaster and emergency situations. Hence, a strong ground exists to explore the opportunity of harnessing social media data for emergency and disaster management in Jakarta (Chatfield & Brajawidagda, 2013). The following section highlights the key success factors of the PetaJakarta.org system.

3. Key success factors of the PetaJakarta.org system

The PetaJakarta.org system has made significant contribution to disaster resilience advancement and there are lessons learnt from its implementation in Jakarta that can be applied to other coastal megacities. The PetaJakarta.org system is one of the first solutions that demonstrate how social media data can be harnessed to generate near real-time city-scale flood maps. Prior to its implementation in Jakarta, the city relied on manually generated flood maps that take approximately six hours to produce and are available in PDF formats only. The PetaJakarta.org system has provided not just a near real-time approach to flood mapping, but also one that is interactive and scalable. During flooding events, the city of Jakarta can now have access to real-time situational information that, on the one hand, allows citizens to safely navigate through the city, and on the other hand, enables government agencies and managing authorities to make more informed decisions relating to flood control and emergency services. For victims of flood disaster, the PetaJakarta.org system is an empowering emergency communication tool because it enables them share information about their immediate situation so that emergency service workers can better locate and support them. Analysis of the PetaJakarta.org dataset (i.e., Twitter conversation) shows that there are several flood reports from Jakarta citizens requesting help from the emergency agency (BPBD), with images attached to depict their situations (See for example, Figure 1). Emergency workers at the control room can view such reports in near real-time, but the extent to which actions are taken to support reporting victims is a discourse of the operational efficiency of an emergency service and is beyond the scope of this study. Overall, citizens' feedback about the usefulness of PetaJakarta.org system was mostly positive, except for very few concerns expressed about issues such as limited geotagged tweets, spamming and geolocation difficulties. For example, a user tweeted: “@petajkt can't do, no specific location can be geotagged on twitter, only main area, so it's useless.” Despite these concerns, PetaJakarta.org online maps were consistently viewed by thousands of Jakarta citizens during major flooding events, suggesting that the system provided a degree of usefulness to citizens. Figure 2 shows Jakarta’s flood situational map generated by the PetaJakarta.org system, with one of the confirmed and retweeted users’ reports highlighted. The key success factors (KSF) of the PetaJakarta.org system can be understood from both the social and technical perspectives as presented below.

Figure 1. Jakarta citizens request emergency help through the PetaJakarta.org system. User identity is obscured to ensure privacy.

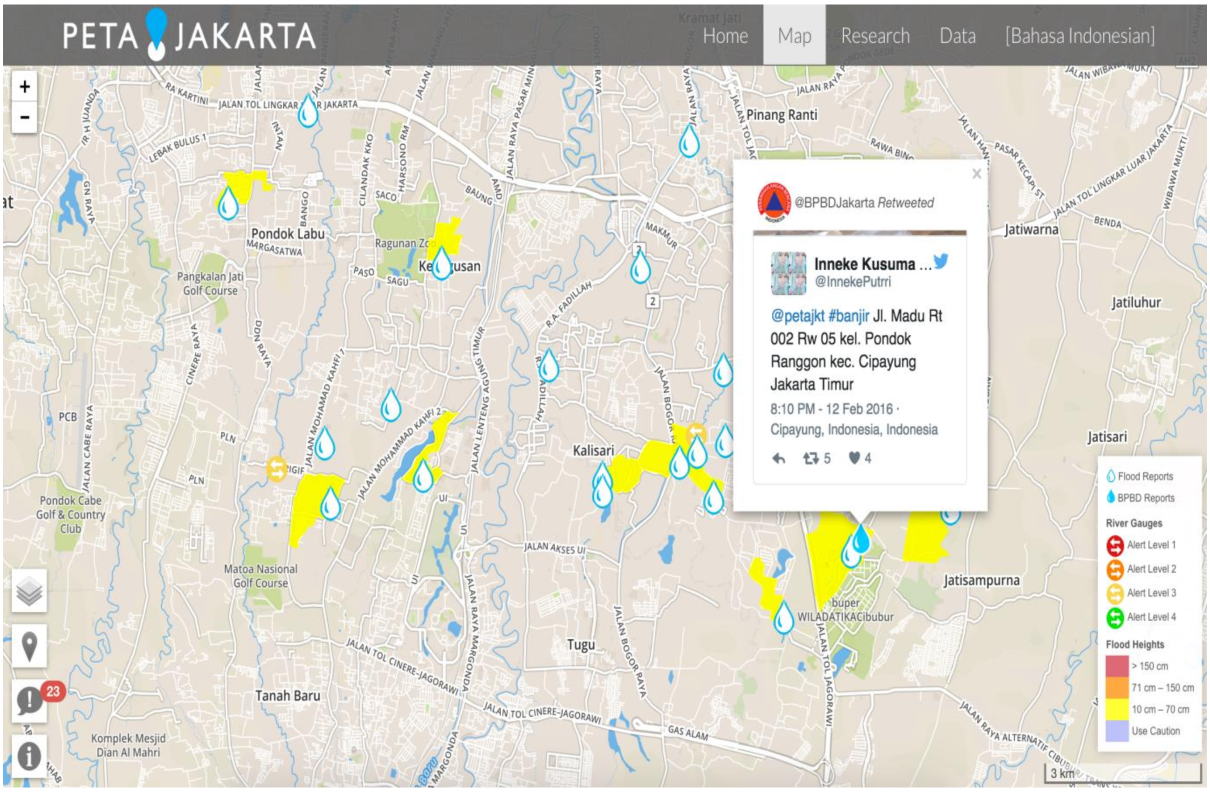

Figure 2. PetaJakarta.org flood situational map

3.1 Social dimension

The PetaJakarta.org project has shown that working collaboratively with authorities is a key success factor for any disaster resilience initiative and has, in many ways, demonstrated how this can be done. The success of PetaJakarta.org depends largely on its collaboration with Twitter Inc., which awarded a Data Grant that allows for unprecedented access to the entire tweet stream during the monsoon season, including historical records of flood-related tweets in Jakarta (Perez et al., 2015). As part of the data grant, there were over 150,000 geolocated tweets contributed by 100,000 users during the 2013/2014 monsoon season. This data revealed both the spatial and temporal characteristics of Jakarta's flood-related tweets and was useful for calibrating the geographical extents of PetaJakarta.org at development phase and for adequately load-testing the underlying software- *CogniCity* prior to real-world deployment. This level of testing, made possible through the data grant from Twitter, allowed the CogniCity software to be successfully deployed for use in Jakarta with minimal hitches during peak demands.

In addition to this data grant support, there was significant expert advice received from the technical team at Twitter during the project implementation in Jakarta. The PetaJakarta.org system also had the rare privilege of being granted Twitter administrative right to automatically send ‘invitation to participate’ messages and reminders to unlimited number of targeted users. Without this administrative right, PetaJakarta.org could not have been as successful in reaching out to Jakarta’s citizens that are tweeting about flooding. The impact of all this invaluable support received from Twitter demonstrates that working with relevant social media authorities enhances the success of crowdsourcing and volunteered geographic information (VGI) initiatives for disaster management.

In addition, collaboration with the BPBD helped the PetaJakarta.org system gain improved acceptance and credibility within the Jakarta government and soon became officially recognised as the digital platform for reporting floods in the city of Jakarta. As part of this collaboration, officials from BPBD and other government agencies both at the Kelurahan (administrative village) and Kecamatan (district) levels were trained to use the PetaJakarta.org platform, with publicity materials provided, including branded mugs and stickers. Community leaders, popularly known in Indonesia as *Lurah* and *Camat* were fully involved in the process. It should be noted that the system development benefited from well-established communication and a cordial relationship between remote techies, local stakeholders in Jakarta and researchers with local, national, and international support units.

Furthermore, eight members of BPBD DKI Jakarta where hosted for nearly two weeks by the SMART Infrastructure Facility, University of Wollongong Australia, where they partook in a series of workshop sessions to discuss BPBD DKI Jakarta’s objectives, workflows, operational structure and any ethnographic concern to be taken into consideration when designing the system. Similarly, in November 2014 and January 2015, BPBD DKI Jakarta hosted groups of students from the University of Wollongong who arrived in Jakarta to collect first-hand data and gain deeper understanding of the various dimensions to the flooding problem in the city. Importantly, PetaJakarta.org maintained staff and an operational office in Jakarta throughout the 3-year period, further allowing for both significant acquisition of local knowledge and avenue to foster strong collaborative relationship through direct support provided to BPBD DKI Jakarta. These types of cross-cultural collaborative sessions are useful and recommended for future applications, especially when the socio-cultural setting in which the system is developed is different from that in which it is to be used.

BPBD DKI Jakarta has recounted the usefulness of the PetaJakarta.org system for gaining flood situational awareness in Jakarta. They acknowledged that the system has helped them reduce the computational time of creating flood maps from a 6-hour manual process to near real time. The PetaJakarta.org system was deployed in the BPBD emergency control room as a source of situational information. The system did not replace but complemented BPBD's traditional disaster management solutions. Despite the benefits of the PetaJakarta.org system to Jakarta emergency service, there were issues identified collectively with BPBD employees directly using the system in the emergency control room. It became obvious during the first year of deployment that there were inadequacies in the system-generated maps when compared with results from physical inspection and manual confirmation of flood situations across the city. For example, the system marked several locations as being flooded; giving an impression that all those locations suffered the same level of flooding, when in fact there were significant differences in the levels of severity. This led to BPBD requesting a change in the system to allow for manual entry by authorised employees in the control room. Hence, by working collaboratively with the emergency team at Jakarta, the *Risk Evaluation Matrix* (REM) interface in the CogniCity software was designed to specifically allow government officials from the emergency service to view, evaluate, and cross-validate data from multiple sources before updating the publicly accessible flood map with information about severity or flood levels at different parts of the city. This kind of co-design constitutes the highest form of participative design- that is, design with (not for) the client.

Media-based publicity is another KSF of the PetaJakarta.org system. The leaders of PetaJakarta.org were quite proactive in promoting the project through media publicity, including via local newspaper articles and radio interviews. Outside Jakarta, the PetaJakarta.org project also had notable press coverage. Reputable organisations such as Amazon, International Federation of Red Cross and Red Crescent Societies, and the United Nations Global Pulse promoted the PetaJakarta.org project through reports on their websites. While the impact of media attention on the overall success of the PetaJakarta.org project could not be measured in a scientifically rigorous manner, it was evident from feedback received from members of the community within and outside Jakarta that the project has gained increased credibility and public attention. For example, in a United States' Federal Communications Commission rule entitled 'Rules Regarding the Emergency Alert System and Wireless Emergency Alerts', PetaJakarta.org was recognised as an example of how a government can leverage crowdsourced data to increase the overall effectiveness of its emergency alert system (Federal Communications Commission, 2016). Similarly, in various international conferences and workshops outside Indonesia, researchers and practitioners in the field of emergency management expressed positive remarks about the PetaJakarta.org project, acknowledging that the project was already known to them through online sources such as news articles, blogs and the project website. Future applications can learn from this rigorous and effective approach to project publicity and promotion of VGI initiatives.

3.2 Technical dimension

From the technical perspective, a note-worthy KSF is PetaJakarta.org's open-source approach and design for system scalability. The PetaJakarta.org project adheres to the principle of free and open access to data and technology. CogniCity, the 'engine room' of the PetaJakarta.org system, is open source software developed using a suite of open source tools. This means that other coastal megacities frequently affected by flood disasters can also implement a civic participatory approach to disaster resilience by freely accessing the CogniCity software (https://github.com/smart-facility/petajakarta-web) and customising it to meet their needs. Free access to this open source software can potentially contribute significantly to disaster resilience in coastal megacities around the world, particularly those in developing nations, where shortage of funding and lack of situational information limit the ability of managing authorities to respond swiftly and effectively to emergency and disaster events (Ogie et al., 2017b). In order to be reusable in different cities and contexts, CogniCity adapted a modular design approach that allows it to scale gracefully to meet changing demands and functionalities. Different cities have different standard operating procedures (SOPs) for managing disaster and emergency events, and the CogniCity software can be customised and deployed to serve these specific contexts. Work is ongoing to fully operationalise the deployment of CogniCity in both Vietnam's Ho Chi Minh City (Saigon) and other flood-prone Indonesian cities such as Jabodetabek (greater Jakarta), Surabaya and Bandung.

Another KSF of the PetaJakarta.org system that presents an interesting positive lesson for future applications is its ability to harvest and integrate data from multiple sources. In addition to data from Twitter, the PetaJakarta.org system was redesigned in 2015 to be able to integrate citizens' flood reports from two other sources, namely *Qlue* and *PasangMata*. Qlue is a civic engagement tool and a map-based platform launched by the government of Jakarta in 2014 to enable citizens report issues of concern such as flood, theft, bribery, traffic, etc. Unlike Twitter that constrains users to 140-character limit (now 280 characters), Qlue accommodates lengthier text messages, with additional requirement that user's reports must always contain an embedded photo or video to help describe the situation. PasangMata is a mobile app that enables citizens to act as news reporters and was launched in 2014 by a Jakarta-based online news agency known as DetikCom. Given that each submitted report is thoroughly verified before being posted on the detikCom news website, flood-related reports acquired from PasangMata are often very reliable. Lengthy text description is also allowed in PasangMata reports.

However, just like Twitter, geolocation of reports and inclusion of photos or videos are optional. This approach of integrating data from multiple sources helps to improve the robustness and reliability of the PetaJakarta.org system. When one platform fails, data sourced from the two other platforms can still keep the PetaJakarta.org system active and running. Subject to availability of submissions from its multiple sources, the PetaJakarta.org system can significantly improve the accuracy of its flood map through cross-platform validation of reports. The PetaJakarta.org website also integrates information about Jakarta's hydrologic and hydraulic infrastructure, including waterways, pumping stations, floodgates, flood gauges and water level sensors. This further contributes to make PetaJakarta.org a reliable and robust system for supporting flood control decision making. Table 1 and Table 2 show a count of reports on the 10 busiest days (by Twitter reports) for each social network used by PetaJakarta.org in 2014/2015 and 2015/2016 monsoon seasons respectively. It should be noted that PetaJakarta.org only harvested geolocated reports from Qlue and PasangMata (referred to as Detik in the table). The 2015/2016 monsoon season recorded fewer flood reports than the 2014/2015 season. A previous study by Ogie et al. (2018b) has already attempted to explain the reason for this drop in flood reports. Though the 2015/2016 season experienced a 29% drop in rainfall, Ogie et al. (2018b) identified other contentious social factors (e.g., users' motivation) that cannot be dismissed without investigating beyond what the project resources has currently allowed. The PetaJakarta.org system records unconfirmed geolocated tweet reports as well as those that have been confirmed by the users (senders) to be related to a real-time flood situation. The percentage of confirmed flood reports is shown for the different 10 days represented in Table 1 and Table 2, with an average of 3.83% for 2014/2015 and 4.37% for 2015/2016.

Table 1. Count of reports and users on the 10 busiest days (by Twitter reports) for each social network used in the 2014/2015 monsoon season.

2014/2015						
Date	Twitter (Non-geolocated)		Twitter (Confirmed geolocated)		Twitter (Unconfirmed geolocated)	Twitter geolocated (Confirmation rate)
	Report	User	Report	User	Report	Ratio of confirmed to unconfirmed reports (%)
20/11/2014	3484	2991	27	16	901	3.00
24/11/2014	1609	1453	20	18	552	3.62
25/11/2014	1612	1443	9	8	376	2.39
27/12/2014	3604	3092	66	33	1180	5.59
23/01/2015	4003	3202	116	28	2414	4.81
01/02/2015	1737	1511	23	17	581	3.96
09/02/2015	18650	16573	346	101	6964	4.97
10/02/2015	14803	13392	233	61	5874	3.97
11/02/2015	5398	4923	72	32	2104	3.42
12/02/2015	2051	1891	19	12	740	2.57

Table 2. Count of reports and users on the 10 busiest days (by Twitter reports) for each social network used in the 2015/2016 monsoon season.

2015/2016								
Date	Twitter (Non-geolocated)		Twitter (Confirmed geolocated)		Twitter (Unconfirmed geolocated)	Twitter geolocated (Confirmation rate)	Qlue	Detik
	Report	User	Report	User	Report	Ratio of confirmed to unconfirmed reports (%)	Report	Report
09/12/2015	473	451	1	1	16	6.25	24	0
30/12/2015	873	769	3	1	70	4.29	63	1
25/01/2016	472	427	2	2	39	5.13	20	2
09/02/2016	756	684	2	1	72	2.78	42	4
12/02/2016	506	419	3	2	49	6.12	46	3
14/02/2016	622	541	1	1	86	1.16	45	5
29/02/2016	545	501	1	1	109	0.92	48	5
25/02/2016	514	461	7	4	111	6.31	63	4
26/02/2016	1718	1518	22	9	278	7.91	141	16
28/02/2016	1059	966	5	5	174	2.87	137	16

Report validation is another KSF of the PetaJakarta.org system. Disaster response is both a safety-critical and time-critical mission. It is therefore imperative that today's socio-technical systems for disaster management do more than simply speed up the aggregation of points on a map. In such systems that aim to harness volunteered geographic information (VGI) for disaster management, it is highly recommended that the quality and accuracy of crowdsourced reports be properly established prior to use (Goodchild, 2007). The PetaJakarta.org system does this well with manual confirmation and expert input as facilitated by its *Risk Evaluation Matrix* (REM).

In this process, reports received from ordinary citizens about flood conditions are treated with less credence and so much effort and time go into verifying the authenticity of such reports before manually updating the flood map with severity information. It is worth noting that this manual confirmation process involves a trade-off between quality/accuracy and time, both of which are critical in disaster response. This issue needs to be addressed in future implementation. Basically, the system needs to be able to identify accurate and high quality social media data, and over time maintain a list of reliable and trustworthy citizens acting as human sensors during flooding events. Previous studies have already shown how to determine the trustworthiness and quality of urban data crowdsourced from multiple sources (Mousa et al., 2015; Prandi et al., 2015). A Platform such as Ushahidi also has an open self-checking mechanism, wherein users from the crowd are enabled to validate reports posted by other users (Gao et al., 2011). However, this approach also has its limitations because it depends on the availability of well-motivated users to serve as report verifiers across various parts of the city and these users may not respond quickly to data verification needs.

4. Challenges of the PetaJakarta system

4.1 Social dimension

A major challenge for the PetaJakarta.org project is weak penetration at community level. PetaJakarta.org did well in media-based publicity and in collaboration with government authorities but failed to engage actively and locally with citizens at the community level, mainly due to limited project resources. PetaJakarta.org campaigns and training sessions focused on government officials, particularly Jakarta’s disaster management agency (BPBD) and this is based on the notion that emergency workers are first responders during crisis events. However, there is a painful lesson that has followed from adopting this approach- a preliminary investigation of participation patterns in the PetaJakarta.org system revealed that both the likelihood of participating (LP) and the likelihood of contributing valuable information (LCVI) were low for ordinary citizens as compared to affiliated users such as emergency workers (Ogie et al., 2018b). Apparently, there are different categories of users, each defined by a different set of participatory capacity based on, for example, motivation, personal goals, skills, access to proper tools and resources, amongst others. A failure to acknowledge and address this participatory inequality undermines the success of the PetaJakarta.org system in harnessing contributions from the millions of citizens who are active on social media and could potentially act as human sensors to support disaster resilience within a megacity context. Future applications should therefore seek to design and implement publicity programmes targeted at all categories of participants, including both affiliated users and ordinary citizens.

User motivation is another major challenge of the PetaJakarta.org system. The PetaJakarta.org system relies on various categories of users to contribute flood-related data, some of whom are affiliated to organisations that directly or indirectly provide rewards for their participations while others are not. For example, participants affiliated to emergency services, local government authorities, and other disaster-specific institutions may easily find motivation to contribute data to the PetaJakarta.org system because they are in an employment relationship with organisations that have vested interest in the success of the data collection initiative. The situation is different for ordinary citizens who are not affiliated to any of such organisations. In many instances, as in the case of the PetaJakarta.org system, it is assumed that natural disasters are communal problems and that relevant data can be harvested opportunistically from citizens without providing them with any personal reward.

However, citizens are strategic social agents, each having a varying set of personal values, needs, goals, norms, and interests that influence how they function within communal and institutional settings (Ogie, 2016). Some citizens are driven by altruism and desire for safer community. Others are more driven by personal interests and desire to profit, knowing fully well that when they engage in capturing and sharing situational information, there are direct and indirect costs associated with the use of network bandwidth, memory, CPU, battery, and personal time. Without rewarding or compensating citizens for the cost incurred in engaging in data collection and sharing, the problem of poor motivation and unwillingness of individuals to participate is likely to undermine the desired response (Kohler et al., 2014). In a previous study, Ogie (2016) highlighted the need for an appropriate incentive to actively engage ordinary citizens in the PetaJarkarta.org system, adding that such reward could simply be social incentives such as community recognition and awards. This issue remains a challenge for the PetaJakarta.org system and needs to be addressed in future implementations.

Another challenge of the PetaJakarta.org system is the lack of long-term funding strategy. A long-term funding strategy that is both achievable and sustainable is crucial for the survival of any free and open data collection initiative. In reality, systems designed to crowdsource disaster-related information require funding to cover operational costs (e.g. web hosting, servers and cloud services, staff salary, publicity and travel cost, incentives for participants, etc.). While the PetaJakarta.org project has been particularly successful in securing short-term funding to facilitate its kick-off and trial over few monsoon seasons, the system was not designed with any strategy to ensure its sustainability in the long run. A means by which this issue can be addressed in future applications is discussed in Section 5.

4.2 Technical dimension

Subjective approach of estimating flood heights is a major challenge of the PetaJakarta.org system. The system relies on social media users acting as human sensors to report flood conditions, including the height of water in their localities. The idea here is that crowdsourced information about flood heights can be used to generate flood maps that show flood level or severity at various parts of the city. The problem with this data collection strategy is that the system aims to retrieve objective data from human opinions, which are inherently subjective. Users' opinion about flood conditions and the way they report that opinion can vary markedly, depending on several factors such as level of education (e.g. ability to estimate flood height in unit of metre), personal motivation (e.g. users' perception of the value of their data to the PetaJakarta.org system), contextual situation (e.g. user's exposure to flood hazards at the time of reporting), system constraints (e.g. Twitter character limit), etc. If humans are to be relied on for objective data, a strategy has to be adopted that minimises errors associated with the subjective nature of human opinions. Clearly, such strategy cannot be based on measurements associated with user's height (e.g. water at knee level) because no standard human height exists, but rather varies from one person to another. A potential solution to this problem is discussed in Section 5.

There is also the challenge of designing for large-scale participation. In the PetaJakarta.org system, social media users who post contents containing the keyword, 'flood' or 'banjir' (in Bahasa Indonesia) are sent an invitation message to confirm that they are reporting about a real-time flood situation and also requesting their participation in contributing to the PetaJakarta.org flood mapping initiative. By design, this message can only be sent once to a user, irrespective of whether he/she opts in or not. Once a user receives this message, his/her username is hashed with an encryption function and stored in a database table containing list of those blocked from further receiving the message in the future. Although, this was done to ensure that users are not overwhelmed with same message each time they use ‘flood’ or ‘banjir’ in their tweets, the approach has negative implication for achieving large-scale participation on the PetaJakarta.org platform. In other words, blocked users, who genuinely want to contribute in the future, are designed out of the system, with little or no information to guide their participation when needed. This is a system design issue that can be avoided in the future by incorporating more robust criteria for inclusion in the list of users blocked from receiving participatory messages. For example, inclusion criteria can be redefined based on the last time a user received a participatory message, the last time a user declined to participate, the user’s frequency of participation, and the user account type (e.g., ordinary citizen or disaster-specific agency that may frequently use the keyword, flood).

The PetaJakarta.org system also faces a few technical challenges that are common to most social media crowdsourcing systems. One of these challenges that are widely accepted as common limitations is semantic difficulty in processing tweets (De Albuquerque et al., 2015; Kusumo et al., 2017; Li et al., 2017). A review of different language schools suggests that a functional grammar (see for example, Halliday, 1973; 1985), not a formal grammar, is a good place to start to address this problem. While formal theories of language view language as rules, functional grammars provide a view of language organised around a principle of use. Ongoing work is being conducted into identifying how functional language theory can be supported computationally.

Another common issue is locational errors arising from using geolocated social media data (Kusumo et al., 2017; Li et al., 2017). When geolocated social media data are used for disaster mapping, it is assumed that users are reporting from the locations where the events being reported have taken place (Kusumo et al., 2017). However, what seems to be a consistent body of evidence in the literature is that disaster-related information on social media tend to come from people who are physically close to the disaster location, not necessarily in the exact disaster location (de Albuquerque et al., 2015; Herfort et al., 2014; Li et al., 2017). This leaves room for locational errors. For example, when the PetaJakarta.org flood reports for 8th March 2016 are superimposed on the BPBD Jakarta official flood map for that day, the result (see Figure 3) shows that there were many flood reports used in the PetaJakarta.org flood mapping system, which were actually outside the areas officially known to have flooded. A previous study which investigated the accuracy of using georeferenced social media data for flood mapping provides several reasons why many of the flood reports used in the PetaJakarta.org system are often outside the areas officially known to have flooded (Ogie & Forehead, 2017). First, people experiencing natural disasters tend to prioritise their safety and take shelter in nearby safer grounds outside the affected areas before engaging in social media to share their experiences. Secondly, in developing nations where internet access is limited, people may be constrained to hold back social media reports of flood observations until they reach an area with reliable internet access. It is also not uncommon for social media users to share second-hand information about flood conditions occurring in a distant location from their neighbourhood, if they are privileged to have received such information through other means. These raise questions about the accuracy of using geolocation metadata to map disasters. A recent study that mapped floods with geolocated social media data minimised such locational errors by using tweets in combination with other relevant authoritative data sources such as digital elevation models (Li et al., 2017).

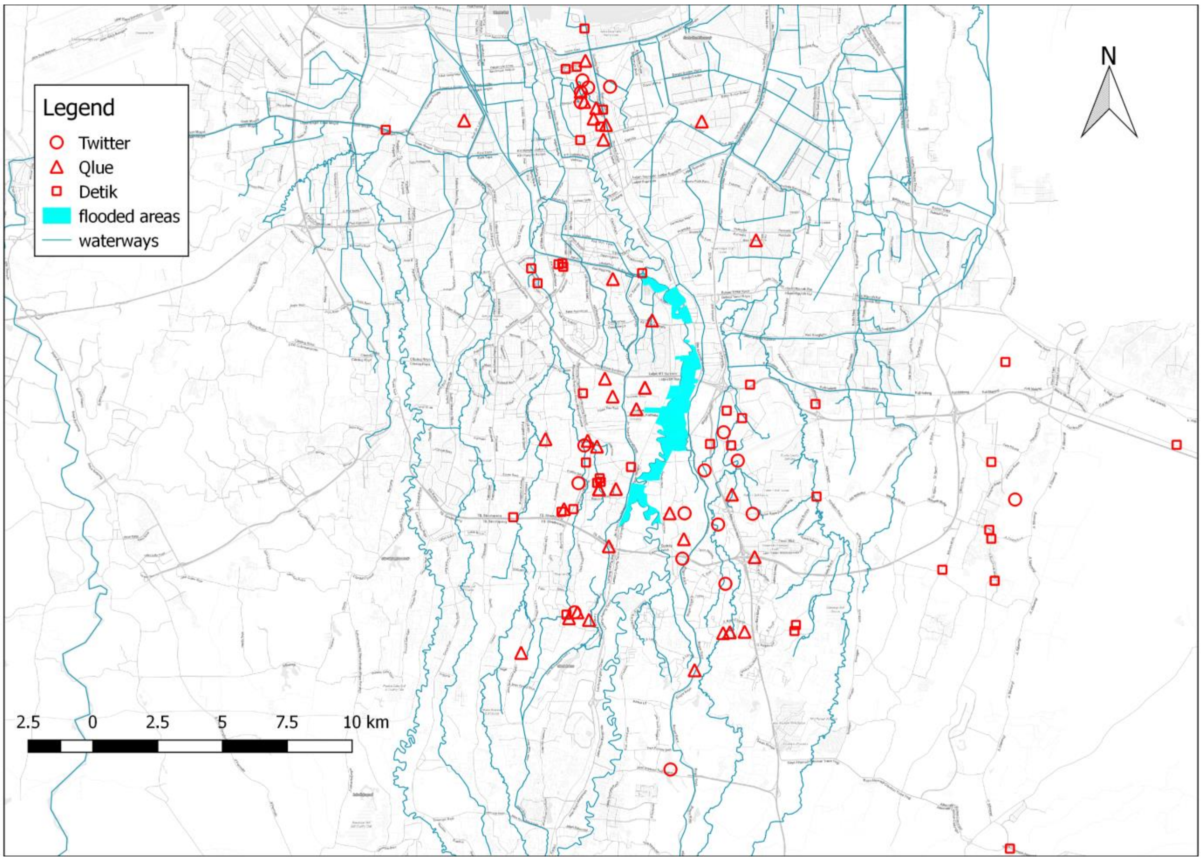

Figure 3. Map comparing PetaJakarta.org flood reports on 8th March 2016 with BPBD official data of flooded areas for same day. Map base layer by Open Street Map/Mapbox

Stamen Toner Light. PetaJakarta.org harvests data from Twitter, Qlue and Detik. Qlue is a government-initiated mapping platform that allows Jakarta's citizens to report various urban issues. Detik operates PasangMata- a social networking app that enables Jakarta's citizens to act as news reporters.

5. Characteristics of future social media crowdsourcing systems

Having highlighted the major challenges experienced in a 3-year implementation of the PetaJakarta.org project, we present some desirable characteristics of future social media crowdsourcing systems that can potentially contribute to address some of the identified challenges. These characteristics are organised into two major perspectives- social and technical dimensions.

5.1 Social characteristics

Certain socially rooted characteristics are desirable to improve future systems used in harnessing VGI and social media data during crisis and disaster events. These characteristics include civic participation, establishment of incentive schemes, and the formulation of a long-term funding strategy to support the operations of the platform. For improved civic participation, all categories of users, including ordinary citizens at local community level should be properly targeted when planning and implementing publicity programmes for civic participation systems used in emergency and disaster situations. PetaJakarta.org has done well in reaching out to government officials and other affiliated users. In future implementations, such level of engagement should be adequately extended to ordinary citizens at local community level, including through briefings in town hall meetings, informational flyers, local press, billboards, local radio and television programmes, community networks, mass mailings and short message service (SMS) dissemination to family households, etc. These civic participatory measures should be planned, not ad hoc.

It is also important that future systems plan for and adopt appropriate incentive mechanisms, which are crucial for attracting and maintaining a list of high quality contributors in any crowdsourcing urban data collection initiative (Ogie, 2016). There are various types of incentive, both monetary and non-monetary (e.g., gamification), that could be adopted to motivate participation in online communities for crowdsourcing volunteered geographic information (VGI) (Bowser et al., 2013; Morschheuser et al., 2016). To improve decision making on adopting an appropriate incentive mechanism, Budhathoki and Haythornthwaite (2013) highlighted several factors that motivate people to participate in open data collection initiatives and they include personal need, learning opportunities, self-efficacy regarding local knowledge, social relationships, personal promotion, altruism, project goal, and monetary reward.

For a project goal that involves crowdsourcing disaster information to support community resilience, monetary incentive (e.g., micropayments) is considered less attractive because it counters altruistic motives behind volunteerism and there are also concerns that the urban poor may overexpose themselves to natural hazards in a bid to earn sufficient money for daily survival, potentially creating ethical and legal risks for responsible authorities (Ogie, 2016). We consider that in such crowdsourcing initiatives that aim to foster urban resilience against natural disasters, incentives that build on altruistic and social values should be used to improve participation. Typically, social incentives such as public recognition, government or community awards/certificates, and non-monetized gifts (e.g., free vouchers to cinemas, museums, zoos, etc.) should be used to encourage civic participation in future systems. It is also worth noting that as volunteers in disaster events tend to be married male adults between the ages of 30 and 50, gamification is becoming a vital means of attracting younger volunteers below the age of 30 (Horita et al., 2014; Meesters et al., 2015).

Furthermore, it is important that future systems used in harnessing VGI and social media data adopt a long-term funding strategy because this is crucial to the viability and sustainability of any free and open data collection initiative (Meijer et al., 2014). Experience from implementing the PetaJakarta.org’s project suggests that future systems will need to source funding to cover operational costs such as web hosting, servers and cloud services, staff salary, incentives for participants, publicity and travel cost, etc. Key to this is a self-funding model that relies on various streams of income such as online donations, advertisements on web interface (e.g., the PetaJakarta.org website), and support from insurance companies that would otherwise suffer financial loss arising from customer claims associated with flood damages.

5.2 Technical characteristics

A number of technically-rooted characteristics are necessary to improve future systems used in harnessing VGI and social media data during crisis and disaster events. These characteristics include minimisation of errors associated with reliance on subjective estimates of flood heights, exploration of novel crisis mapping alternatives to social media-based geolocation metadata, and the adoption of a carefully designed data crowdsourcing strategy. One major concern with the PetaJakarta.org flood mapping system is that it relies on humans to arbitrarily provide estimates of flood heights in their localities. To minimise errors associated with the subjective nature of human opinions, it is important that future applications adopt a strategy that is based on a combination of several factors.

Basically, there needs to be a significant number of citizen-generated reports of flood height in a given area, which together provide a consistent and reliable body of knowledge for accurately inferring the flood extent in that area. The process needs to be guided by factors such as (1) expected range of flood height values within the given area as derived from historical records of flooding events, (2) weight-based averaging of flood heights, where submissions from users who are historically known to be trustworthy and accurate in their estimations are given higher weights, and (3) restricting use of citizens' reports for a given area to only those received within the same defined time window. The above strategy is obviously a more rigorous approach for using citizens' reports to estimate flood heights or severity of inundation, but one that should prove more rewarding in terms of the accuracy and reliability of the co-produced flood map.

The accuracy and reliability of crowdsourced disaster mapping can be further improved by reconsidering the basis for determining locations of events. As earlier discussed in Section 4.2, the use of geolocation metadata as a basis for disaster or crisis mapping can potentially produce inaccurate results. Future systems can address this issue by adopting emerging Artificial Intelligence technology for geolocation based on photo pixels as demonstrated in Google's PlaNet (Weyand et al., 2016).

Accuracy remains a crucial issue in the context of disaster response and future initiatives involving crowdsourced urban data collection can follow one of two strategies to minimise errors in harnessing VGI and social media data. The first strategy is based on opportunistic harvesting of disaster-related information posted on social media sites (Ogie, 2016). This approach is advantageous from the perspective that it allows a vast amount of data to be harnessed freely without bothering about incentive mechanisms for users. However, the quality and usefulness of data collected through this approach is questionable, particularly for the purpose of supporting emergency and disaster management. The reason being that social media contents harvested opportunistically are often very noisy, unstructured, and unfit for purpose (Sui and Goodchild, 2011). To derive utility from this type of data, the system must be robust enough and carefully designed with high-level accuracy in detecting, filtering, and understanding social media contents. Palen and Anderson (2016) also recommend capturing and analysing contextual information, including historical online behaviours of users, to ensure the retrieved social media content is fit for purpose. This is obviously a challenging task to accomplish computationally and the PetaJakarta.org system has not managed to do so.

The second strategy is based on participatory harvesting of disaster-related information, made possible not through existing social media sites, but through third-party applications designed purposefully to crowdsource structured data from citizens. The advantage of this approach is that information is requested in specific formats and the contents received are often defined for purpose, so that the filtering or data cleansing required is minimal. However, the downside is that the uptake and sustainability of systems designed based on participatory approach depend largely on appropriate incentive mechanisms for rewarding or motivating participants (Ogie, 2016). The PetaJakarta.org system clearly needs to improve its harvesting strategy. In principle, the PetaJakarta.org system harvests data opportunistically, but does so with a data filtering algorithm that is not robust enough to handle the noise and the linguistic complexities within crowdsourced social media contents. To make up for these lapses, PetaJakarta.org embarked on data collection campaigns that expressly request users to contribute data purposely and objectively, but again this strategy was adopted without a proper incentive mechanism required to drive participatory harvesting of urban data.

7. Conclusion

The application of crowdsourced social media data in disaster management is a growing field of research, which needs to benefit more from real-world experiments. This paper, which reports on lessons learnt from real-world implementation of a highly-praised project (PetaJakarta.org), contributes in a number of ways to the literature and the practice of applying crowdsourced social media data in emergency and disaster events. The paper started with a brief background of the PetaJakarta.org project that involves crowdsourcing flood-related data from citizens of Jakarta. The paper then highlighted key success factors of the PetaJakarta.org system that can be adopted as strategies to enhance future applications. These factors include working collaboratively with authorities, adopting media-based publicity to improve awareness, incorporating the ability to harvest data from multiple streams, and designing for scalability and reusability.

Several challenges of the PetaJakarta.org system, which should be avoided in future applications, were also discussed. These challenges include weak penetration at community level, motivational issues, lack of long-term funding strategy, and subjective approach to estimating objective data. Furthermore, the study contributes by describing desirable characteristics of future social media crowdsourcing systems, which when adopted can help to address some of the identified challenges. It is hoped that the knowledge and lessons shared in this paper will motivate and enable other coastal megacities to improve how they apply crowdsourced social media data in disaster management, particularly in the context of using free and customisable open source software such as CogniCity.

At the time of this study, the PetaJakarta project has been replaced by *PetaBencana.* PetaBencana builds on the relationships, knowledge, software architecture, and strengths of the PetaJakarta project, taking on board its key success factors while striving to improve through lessons from its shortcomings. Significant effort has been made with PetaBencana to minimise errors in the crowdsourced flood maps. Unlike PetaJakarta that focused mainly on Jakarta city, PetaBencana extends beyond Jakarta to include other Indonesian cities such as Surabaya, Bandung, and Semarang.

Acknowledgements

This work was supported by the Australian National Data Service (ANDS) through the National Collaborative Research Infrastructure Strategy Program [ANDS MODC 15, 2014], the Department of Foreign Affairs and Trade, Australia (DFAT) [agreement number 71984] and the University of Wollongong Global Challenges Program.

References

Abel, F., Hauff, C., Houben, G.J., Stronkman, R. and Tao, K. (2012), April. Twitcident: fighting fire with information from social web streams. In *Proceedings of the 21st International Conference on World Wide Web* (pp. 305-308). ACM.

Acar, A. and Muraki, Y. (2011). Twitter for crisis communication: lessons learned from Japan's tsunami disaster. *International Journal of Web Based Communities, 7*(3), pp.392-402.

Bowser, A., Hansen, D., He, Y., Boston, C., Reid, M., Gunnell, L. and Preece, J. (2013), October. Using gamification to inspire new citizen science volunteers. In *Proceedings of the first international conference on gameful design, research, and applicatio*ns (pp. 18-25). ACM.

Brandusescu, A. and Sieber, R.E. (2017), The spatial knowledge politics of crisis mapping for community development. *GeoJournal*, pp.1-16. DOI 10.1007/s10708-017-9784-9

Budhathoki, N.R. and Haythornthwaite, C., 2013. Motivation for open collaboration: Crowd and community models and the case of OpenStreetMap. *American Behavioral Scientist, 57*(5), pp.548-575.

Carley, K.M., Malik, M.M., Kowalchuck, M., Pfeffer, J. and Landwehr, P., 2015. Twitter Usage in Indonesia. Technical report, Institute for Software Research, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213

Chatfield, A.T. and Brajawidagda, U., 2013, January. Twitter early tsunami warning system: A case study in indonesia's natural disaster management. In *System sciences (HICSS), 2013 46th Hawaii international conference on* (pp. 2050-2060). IEEE.

Coiera, E., 2007. Putting the technical back into socio-technical systems research. *International journal of medical informatics*, *76*, pp.S98-S103.

Croitoru, A., Crooks, A., Radzikowski, J. and Stefanidis, A., 2013. Geosocial gauge: a system prototype for knowledge discovery from social media. *International Journal of Geographical Information Science, 27*(12), pp.2483-2508.

Crooks, A.T. and Wise, S., 2013. GIS and agent-based models for humanitarian assistance. *Computers, Environment and Urban Systems*, *41*(2013), pp.100-111.

De Albuquerque, J.P., Herfort, B., Brenning, A. and 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), pp.667-689.

Dennis, J.B., 2017. Principles to Support Modular Software Construction. *Journal of Computer Science and Technology*, *32*(1), pp.3-10.

Eilander, D., Trambauer, P., Wagemaker, J. and van Loenen, A., 2016. Harvesting social media for generation of near real-time flood maps. *Procedia Engineering*, *154*, pp.176-183. https://doi.org/10.1016/j.proeng.2016.07.441.

Federal Communications Commission, 2016. Rules Regarding the Emergency Alert System In: Report and Order and Further Notice of Proposed Rulemaking, FCC 16-127, Before the Federal Communications Commission Washington, D.C. 20554. https://docs.fcc.gov/public/attachments/FCC-16-127A1.pdf.

Gao, H., Barbier, G. and Goolsby, R., 2011. Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems*, *26*(3), pp.10-14.

Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal, 69*(4), pp.211-221.

Granell, C. and Ostermann, F.O., 2016. Beyond data collection: Objectives and methods of research using VGI and geo-social media for disaster management. *Computers, environment and urban systems, 59*(2016), pp.231-243.

Halliday, M. A. K. (1985) *An Introduction to Functional Grammar* London: Edward Arnold

Halliday, M. A. K. (1973) “The Functional Basis of Language” in Webster, J. J. ed/ (2003) *M. A. K. Halliday Collected Works Volume 3: On Language and Linguistics,* Chapter 14, p. 298-320; First published in Bernstein, B. ed/ (1973) *Applied Studies towards a Sociology of Language, Vol. 2: Class Codes and Control* London Routledge and Kegan Paul, pp. 343-366

Herfort, B., de Albuquerque, J.P., Schelhorn, S.J. and Zipf, A., 2014. Exploring the geographical relations between social media and flood phenomena to improve situational awareness In J. Huerta, S. Schade, & C. Granell (Eds.). *Connecting a digital Europe through location and place* (pp. 55–71). Cham, Switzerland: Springer International. doi: 10.1007/978-3-319-03611-3_4

Holderness, T. and Turpin, E., 2015. From Social Media to GeoSocial Intelligence: Crowdsourcing Civic Co-management for Flood Response in Jakarta, Indonesia. In *Social media for government services* (pp. 115-133). Springer International Publishing.

Horita, F.E., Assis, L.F.F.G., Castanhari, R.E., Isotani, S., Cruz, W.M. and de Albuquerque, J.P., 2014. A Gamification-based Social Collaborative Architecture to increase resilience against natural disasters. *Simpósio Brasileiro de Sistemas de Informação (SBSI), Londrina, Brazil.*

Huang, C. M., Chan, E., & Hyder, A. A. (2010). Web 2.0 and internet social networking: A new tool for disaster management?—Lessons from Taiwan. *BMC Medical Informatics and Decision Making, 10*(57), 10–57.

Kohler, T., Steghöfer, J.P., Busquets, D. and Pitt, J. (2014). The value of fairness: trade-offs in repeated dynamic resource allocation. In *Self-Adaptive and Self-Organizing Systems (SASO), 2014 IEEE Eighth International Conference on* (pp. 1-10). IEEE.

Kumar, S., Barbier, G., Abbasi, M.A. and Liu, H., 2011, July. TweetTracker: An Analysis Tool for Humanitarian and Disaster Relief. In *Fifth International AAAI Conference on Weblogs and Social Media*. Barcelona Spain:
AAAI.

Kusumo, A.N.L., Reckien, D. and Verplanke, J., 2017. Utilising volunteered geographic information to assess resident's flood evacuation shelters. Case study: Jakarta. *Applied Geography, 88*, pp.174-185.

Li, Z., Wang, C., Emrich, C.T. and Guo, D., 2017. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. *Cartography and Geographic Information Science*, pp.1-14.

Lu, C.C., Ying, K.C. and Chen, H.J., 2016. Real-time relief distribution in the aftermath of disasters–A rolling horizon approach. *Transportation research part E: logistics and transportation review*, *93*, pp.1-20.

Meesters, K., Ruhe, A. and Soetanto, M., 2015, October. Gamification for data gathering in emergency response exercises. In *European Conference on Games Based Learning* (p. 672). Academic Conferences International Limited.

Meijer, A., de Hoog, J., van Twist, M., van der Steen, M. and Scherpenisse, J., 2014. Understanding the dynamics of open data: From sweeping statements to complex contextual interactions. In *Open government* (pp. 101-114). Springer New York.

Morschheuser, B., Hamari, J. and Koivisto, J., 2016, January. Gamification in crowdsourcing: a review. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on* (pp. 4375-4384). IEEE.

Mousa, H., Mokhtar, S.B., Hasan, O., Younes, O., Hadhoud, M. and Brunie, L., 2015. Trust management and reputation systems in mobile participatory sensing applications: A survey. *Computer Networks*, *90*, pp.49-73.

Obe, R.O. and Hsu, L.S., 2015. *PostGIS in action*. Manning Publications Co.

Ogie, R.I., 2016. Adopting incentive mechanisms for large-scale participation in mobile crowdsensing: from literature review to a conceptual framework. *Human-centric Computing and Information Sciences*, *6*(1), p.24.

Ogie, R.I. and Forehead, H., 2017, December. Investigating the accuracy of georeferenced social media data for flood mapping: The PetaJakarta. org case study. In *Information and Communication Technologies for Disaster Management (ICT-DM), 2017 4th International Conference on* (pp. 1-6). IEEE.

Ogie, R.I., Holderness, T., Dunn, S. and Turpin, E., 2016. Vulnerability analysis of hydrological infrastructure to flooding in coastal cities-a graph theory approach. Transforming the future of infrastructure through smarter information: *Proceedings of the international conference on smart infrastructure and construction (ICSIC)*, 27–29 june 2016, Cambridge, UK (2016), pp. 633–638

Ogie, R.I., Holderness, T., Dunn, S. and Turpin, E., 2018a. Assessing the vulnerability of hydrological infrastructure to flood damage in coastal cities of developing nations. *Computers, Environment and Urban Systems, 68*(2018), pp. 97-109

Ogie, R.I., Dunn, S., Holderness, T. and Turpin, E., 2017a. Assessing the vulnerability of pumping stations to trash blockage in coastal mega-cities of developing nations. *Sustainable Cities and Society*, *28*, pp.53-66.

Ogie, R.I., Forehead, H., Clarke, R.J. and Perez, P., 2018b. Participation patterns and reliability of human sensing in crowd-sourced disaster management. *Information Systems Frontiers*. DOI https://doi.org/10.1007/s10796-017-9790-y

Ogie, R., Holderness, T., Dunbar, M. and Turpin, E., 2017b. Spatio-topological network analysis of hydrological infrastructure as a decision support tool for flood mitigation in coastal mega-cities. *Environment and Planning B: Urban Analytics and City Science*, *44*(4), pp.718-739.

Palen, L. and Anderson, K.M., 2016. Crisis informatics—New data for extraordinary times. *Science, 353*(6296), pp.224-225.

Perez, P., du Chemin, T.H., Turpin, E. and Clarke, R., 2015, September. Citizen-driven flood mapping in Jakarta: a self-organising socio-technical system. In *Self-Adaptive and Self-Organizing Systems Workshops (SASOW), 2015 IEEE International Conference on* (pp. 174-178). IEEE.

Prandi, C., Ferretti, S., Mirri, S. and Salomoni, P., 2015, March. Trustworthiness in crowd-sensed and sourced georeferenced data. In *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on* (pp. 402-407). IEEE.

Roche, S., Propeck-Zimmermann, E. and Mericskay, B., 2013. GeoWeb and crisis management: Issues and perspectives of volunteered geographic information. *GeoJournal, 78*(1), pp.21-40.

Rogstadius, J., Vukovic, M., Teixeira, C.A., Kostakos, V., Karapanos, E. and Laredo, J.A., 2013. CrisisTracker: Crowdsourced social media curation for disaster awareness. *IBM Journal of Research and Development*, *57*(5), pp.4-1.

Shanley, L., Burns, R., Bastian, Z. and Robson, E., 2013. Tweeting up a storm: the promise and perils of crisis mapping, *Photogrammetric Engineering and Remote Sensing*, *79*(10), pp.865–879.

Sui, D. and Goodchild, M., 2011. The convergence of GIS and social media: challenges for GIScience. *International Journal of Geographical Information Science*, *25*(11), pp.1737-1748.

Tsai, J.L., 2009. Convertible multi-authenticated encryption scheme with one-way hash function. *Computer Communications, 32*(5), pp.783-786.

Weyand, T., Kostrikov, I. and Philbin, J., 2016, October. Planet-photo geolocation with convolutional neural networks. In *European Conference on Computer Vision* (pp. 37-55). Springer International Publishing.

Yates, D. and Paquette, S., 2011. Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. *International journal of information management*, *31*(1), pp.6-13.

Zook, M., Graham, M., Shelton, T. and Gorman, S., 2010. Volunteered geographic information and crowdsourcing disaster relief: a case study of the Haitian earthquake. *World Medical & Health Policy, 2*(2), pp.7-33.