IT Enabled Crowds: Leveraging the Geomobile Revolution for Disaster Management

Marta Poblet[1,2], Esteban García-Cuesta,[3] Pompeu Casanovas[1]

RMIT University, Melbourne VIC 3001 Australia
[1]UAB Institute of Law and Technology, Campus UAB, 08193 Bellaterra, Spain
[3]iSOCO, Madrid, Spain

marta.pobletbalcell@rmit.edu.au, egarcia@isoco.com, pompeu.casanovas@uab.cat

Abstract. This paper offers an exploratory approach to crowdsourcing methods, tools, and roles based on different levels of involvement of users, skills required, and types of data being processed (from raw data to highly structured dada). The paper also aims at refining different crowdsourcing categories and opening up a theoretical discussion on the advantages and limits of using crowdsourcing methods and technologies in disaster management activities.

Keywords: crowdsourcing, crowdsensing, micro-tasking, data, online platforms, disaster management

1 Introduction

Mobile technologies, location-based services, and geospatial data are currently fuelling the geomobile revolution that brings up to the front the relational dimension of space. A growing plethora of sensors and applications in our mobile devices are constantly producing data, both for us and about us: geospatial coordinates in digital maps, routes, check-ins, etc. Such geospatial data are the digital anchors from which we interact with our immediate context. These data also contribute to augment our reality with added layers of information. Likewise, our real-time geopositioning in a given space triggers the digital footprints that we leave as we interact with our immediate context.

Combined with different crowdsourcing approaches and methods, the geomobile revolution also creates unprecedented opportunities for research, industry, and social change. This paper explores how state-of-the-art technologies are opening up new avenues for citizens' involvement in disaster management initiatives with different crowdsourcing roles.

2 Crowdsourcing: the power of the crowds

The term crowdsourcing was first coined by Jeff Howe in 2006 when referring to "the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call" [11]. Since Howe's first definition, different crowdsourcing categories, dimensions, and typologies have been discussed in the literature [5,18,7,6,19,8,17]. Other studies consider crowdsourcing as part of the broader paradigm of collective intelligence [12] and review the similarities, overlapping and gaps between human computation, crowdsourcing, social computing and data mining [16].

The three key elements intersecting in Web-based crowdsourcing are the crowd, the outsourcing model, and advanced Internet technologies [17]. According to their definition, "crowdsourcing is a sourcing model in which organizations use predominantly advanced Internet technologies to harness the efforts of a virtual crowd to perform specific organizational tasks" [17]. Another recent definition by Chamales also highlights the technological component of crowdsourcing [3]:

Crowdsourcing technology brings together a distributed workforce of individuals in order to collect resources, process information, or create new content. The implementation of a crowdsourcing system can vary widely, from complex online websites that coordinate a million simultaneous workers to low-tech, ad hoc approaches that use a shared spreadsheet." [3].

At present, Web 2.0 technologies have expanded the range of available crowdsourcing methods to the point that the concept has become an umbrella term that covers multiple ways to collect and share information, respond to labor offers or contests, volunteer for a number of tasks, etc. Reviewing some of the currently available tools will provide the basis for some useful distinctions.

3 Open source crowdsourcing platforms

In the last few years, crowdsourcing platforms have sprouted to leverage the resources of the crowds in crisis and disaster management efforts [14]. Most of these tools have embraced open source licenses from their inception. The first generation of open-source platforms, Ushahidi, OpenStreetMap, and Sahana are among the most popular, with large communities of developers and users. Ushahidi was initially launched as a Google Maps mash-up to map reports of violence after the Kenyan post-election fallout at the beginning of 2008.[1] Ushahidi and Crowdmap (its hosted version) have been used in over 30,000 deployments in 156 countries [9].

[1] http://www.ushahidi.com/about-us

OpenStreetMap is an editable map with more than 1,350,000 registered users (as of August 2013).[2] The platform, started in 2004 by Steve Coast, allows free access to the full map dataset via the Open Data Commons Open Database License (ODbL).

The origins of Sahana ("relief" in Sinhala) are also grounded in the response to a critical event (the Indian Ocean Tsunami of 2004) and the need to coordinate organizational efforts. The newer version of the platform (Eden) specifically addresses disaster management tasks and includes dedicated modules for organization registry, volunteer management, and online mapping.[3]

CrisisTracker, initially developed by Jakob Rogstadius in 2011, combines automated processing with crowdsourcing to quickly detect new events in Twitter. The CrisisTracker platform uses an automated real-time clustering algorithm based on Locality Sensitive Hashing (LSH) to group together tweets that are textually very similar.[4] Volunteers are then invited to refine the topical clusters or create new ones.

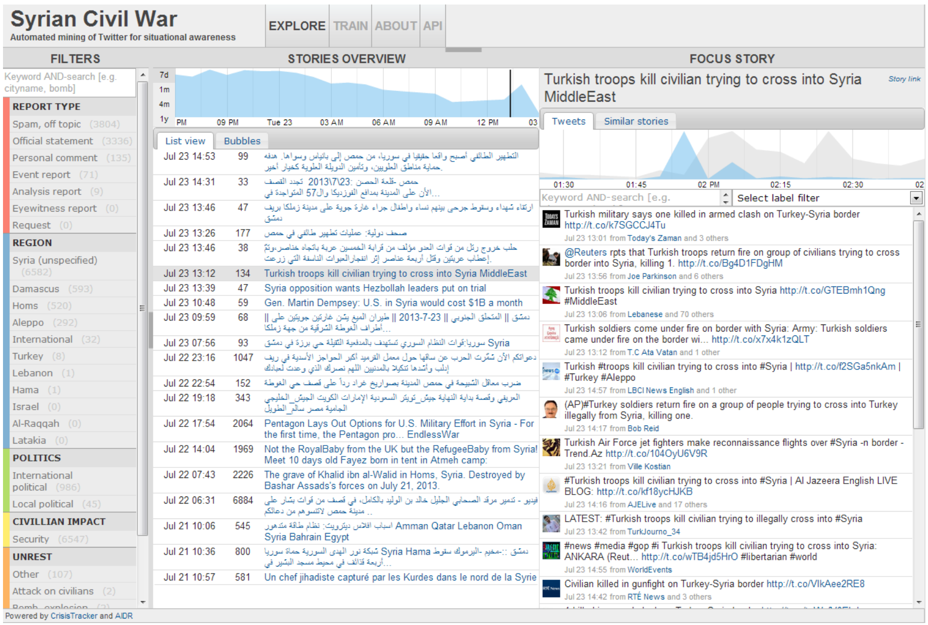

Fig. 1. Clustering tweets related to the Syrian Civil War with CrisisTracker

Deployments and projects using any of the above platforms typically require help from volunteers with different skill levels in information and data management (i.e. media monitoring, categorization, reporting, etc.), GIS (geolocation and mapping) or disaster management (logistics, volunteer management, etc.). In recent years, a number of volunteer communities from different backgrounds and domains have gathered

2 https://openstreetmap.org
3 https://eden.sahanafoundation.org/
4 https://github.com/jakobrogstadius/crisistracker

around these tools to contribute to crisis and disaster management efforts.[5] Frequently, volunteer communities provide initial training to any incoming contributor via different channels: skype chats, IRC channels, Google hangouts, tutorials, etc.

A second generation of open-source tools mainly consists of lightweight, easy-to-navigate mobile applications. In this mobile environment, the tasks (or micro-tasks) that volunteers are usually requested to complete are bite-size chunks (e.g. translate a sentence, tag a tweet, assess different levels of damage as seen in a picture, etc.). TaskMeUp is an application initially developed in 2010 by InSTEDD.org where users can request the help of volunteers on tasks such as text message translation or categorizing information.[6] Crowdcrafting, defines itself as a “framework for developing and deploying crowd-sourcing and microtasking apps”.[7] Recently, two of these microtasking apps—TweetClicker and ImageClicker—have been used as part of the response to typhoon Laura in the Philippines in a partnership between UNOCHA and digital volunteer organizations (i.e. the Standby Task Force and Humanitarian Open Street Map). The tasks requested to volunteers—in an open call open to the general public—were fairly simple. TweetClicker asked to tag a tweet at a time (from a set of tweets filtered with machine learning algorithms) either as not relevant to the disaster, as a request for help, as reporting infrastructure damage, or a population displacement. Similarly, ImageClicker proposed three categories of damage (none, mild, or severe) to tag images extracted from social media. Each app included a mini-tutorial to guide volunteers, who could also participate in a skype chat if they wanted to share questions or comments. The two apps have been developed by Micromappers, a project led by Patrick Meier at the Qatar Computing Research Institute (QCRI).[8]

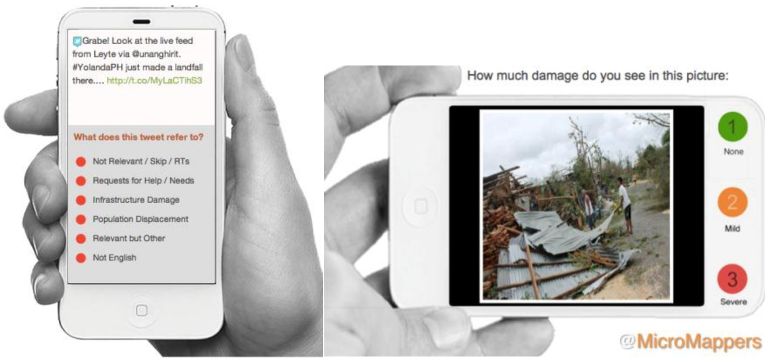

Fig. 2. TweetClicker and ImageClicker (by MicroMappers)

---

5 See the Digital Humanitarian Network, a network of volunteer organizations working on disaster management efforts from different backgrounds, http://digitalhumanitarians.com/

6 https://bitbucket.org/instedd/taskmeup

7 http://crowdcrafting.org/about

8 http://www.qcri.com/

4 Crowdsourcing roles

The size and composition of the crowd can also help to determinate whether the crowdsourced effort is unbounded (anyone can participate) or bounded to "a small number of trusted individuals" [13]. We can further distinguish the role of the crowd based on the type of data being processed and the level of participation involved. This leads to four types of crowdsourcing roles based on: (i) type of data processed (raw, semi-structured, and structured data), (ii) participants' level of involvement (passive or active) and, (iii) skills required to fulfill the assigned task (basic or specialized skills). Figure 3 below shows these four roles based on how the crowd is involved in the process of generating and adding value to the knowledge chain process.

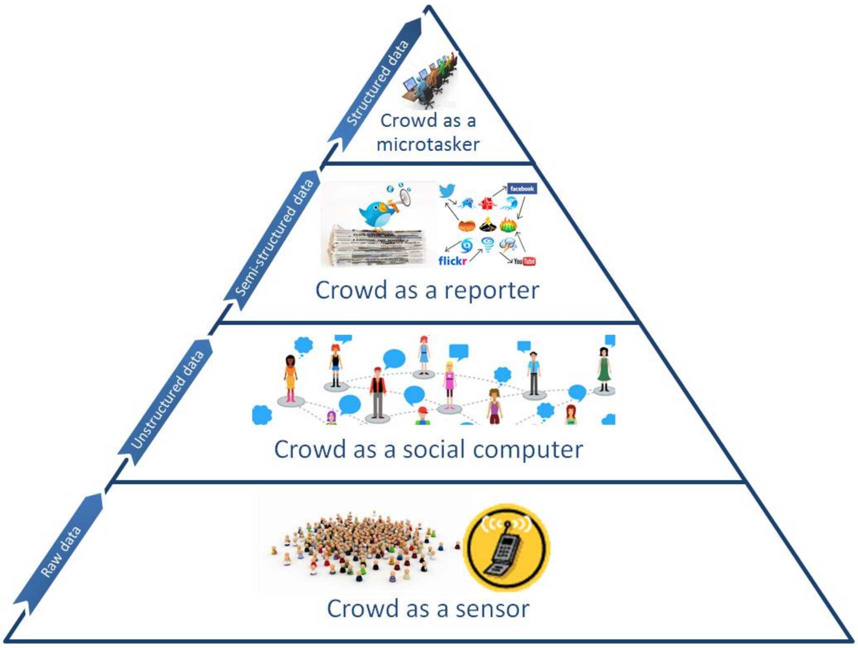

Fig. 3. Crowdsourcing roles based on users' involvement and level of data processing

The lower tiers of the pyramid represent users who generate raw or unstructured data by the mere use of mobile phones, tablets, etc. (crowd as a sensor) or their either occasional or regular use of social media (crowd as a social computer). In contrast, the two top tiers include users with an explicit, conscious use of a priori knowledge to achieve a specific goal (crowd as a reporter and crowd as a microtasker). Moving from lower to higher levels in the pyramid also implies a shift in the quality of the obtained data. From a knowledge generation and data processing point of view we are ranging from raw data, unstructured data, or semi-structured data, to structured data (which also become interpreted data resulting from the execution of the process).

Whereof, lower roles in the pyramid produce raw data and higher roles high valued data which are related with the action of solving a specific problem (e.g. labeling an image). Such a categorization also implies different levels of effort by the crowds involved:

i) Crowd as sensors: people generate raw data just because some processes are automatically performed by sensor-enabled mobile devices (e.g. processes run in the backend by GIS receivers, accelerometers, gyroscopes, magnetometers, etc.) which can be later on used for a purpose (i.e. mobile phone coordinates for positional triangulation, traffic flow estimates, etc.). This type of data collection has been defined elsewhere as “opportunistic crowdsourcing” [30]. Opportunistic crowdsourcing requires very low data processing capabilities (if any) on the side of participants and is the most passive role in the contributing information chain.

ii) Crowd as social computers: people generate unstructured data mostly by using social media platforms for their own communication purposes (e.g. sharing contents or socializing in Facebook, Twitter, Instagram, etc.). Social media users do not process information in any specific form, but these data can later be reused to extract semantically structured information. As in the previous role, there is no explicit participatory effort in any crowdsourced initiative or project).

iii) Crowd as reporters: people offer first-hand, real-time information on events as they are unfolding (e.g. they tweet about a hurricane making landfall and the reporting damages in a specific location). This user-generated content already contains valuable metadata added by users themselves (e.g. hashtags) than can be used as semi-structured, preprocessed data.

iv) Crowd as microtaskers: people generate structured, high quality, interpreted data by performing some specific tasks over raw data (e.g. labeling images, adding coordinates, tagging reports with categories, etc.). This role requires an active participation of users in the crowdsourcing effort and it may exploit special skills or require different levels of previous training.

5 Conclusion

As new tools and technologies enable citizens to participate in crowdsourced efforts with different roles and skills, new opportunities emerge for projects and initiatives involving the management of large amounts of data. Disaster management, environmental sensing, scientific research, business, and marketing are among the areas than can benefit from crowdsourced input or microtasking activities. The efficient allocation of tasks to a largely distributed online workforce can produce immediate outcomes that would be otherwise difficult to obtain with traditional outsourcing methods. Yet, bringing such a large crowd-force into organizational workflows raises a number of issues that need to be taken into account: management of the crowd, accuracy, reliability, quality control of the outcomes, etc. Compliance with data protection and privacy rules (including different types of consent) will also help to delimitate how crowdsourced data can be aggregated, shared, used, and reused. Finally, an appropriate ethical framework can be brought into the picture to interact and complement rules, principles, and standards whenever needed [1,2,15]. Further research will explore further the connections between different crowdsourcing roles and the corresponding regulatory frameworks.

6 Acknowledgements

This research has been supported by two research grants from the Spanish Ministry of Economy and Competitiveness (MINECO) to the projects “CrowdCrissControl” (IPT-2012-0968-390000) and “Crowsourcing: instrumentos semánticos para el desarrollo de la participación y la mediación online” (DER 2012- 39492 -C02 -01).

7 References

1. Casanovas, P., Poblet, M. 2008. Concepts and Fields of Relational Justice. In: Sartor, G., Casanovas, P., Casellas, N., Rubino, R. (Eds) Computable Models of the Law: Languages, Dialogue, Games, Ontologies. LNAI 4884, Springer Verlag, Berlin, Heidelberg: 323-339.
2. Casanovas, P., Poblet, M. 2009. The Future of Law: Relational Law and next Generation of Web Services. In: M. Fernández-Barrera et al. (Eds), The Future of Law and Technology: Looking into the Future. Selected Essays. European Press Academic Publishing, Florence: 137-156.
3. Chamales, G. 2013. Towards trustworthy social media and crowdsourcing. Wilson Center Commons Lab. Available at http://www.wilsoncenter.org/sites/default/files/TowardsTrustworthySocialMedia_FINAL.pdf.
4. Chatzimilioudis, G., Konstantinidis, A., Laoudias, C., Zeinalipour-Yazti, D. 2012. Crowdsourcing with smartphones. Internet Computing, IEEE, 16(5), 36-44.
5. Doan, A., Ramakrishnan, R., and Halevy, A. Y. 2011. Crowdsourcing systems on the World-Wide Web, *Communications of the ACM,* Vol. 54 (4) 86-96.
6. Estellés-Arolas, E., González-Ladrón-de-Guevara, F. 2012. Towards an integrated crowdsourcing definition. *Journal of Information Science*, Vol. 38 (2): 189–200.
7. Geiger, D., Seedorf, S., Schulze, T., Nickerson, R. C., Schader, M. 2011. Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes. AMCIS-Proceedings of the Seventeenth Americas Conference on Information Systems (Detroit, Michigan August 4th-7th). Available at http://schader.bwl.uni-mannheim.de/fileadmin/files/schader/files/publikationen/Geiger_et_al._-_2011_-_Managing_the_Crowd_Towards_a_Taxonomy_of_Crowdsourcing_Processes.pdf.
8. Haklay, M. 2013. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In *Crowdsourcing Geographic Science,* D. Sui, S. Elwoold, M. Goodchild (eds.): 105-122, Springer Netherlands.

9. Hersman, E. 2012. Which is better, 10,000 reports or 10,000 maps? (Part 1/2), http://blog.ushahidi.com/2012/07/10/which-is-better-10000-reports-or-10000-maps-part-12/
10. Hetmank, L. 2013. Components and Functions of Crowdsourcing Systems: A Systematic Literature Review. *Wirtschaftsinformatik Proceedings 2013.* Paper 4.
11. Howe, J. 2006. The rise of crowdsourcing. *Wired*, June 14 2006. Available at http://www.wired.com/wired/archive/14.06/crowds.html
12. Malone, T. W., Laubacher, R., Dellarocas, C. N. 2009. Harnessing crowds: Mapping the genome of collective intelligence. MIT Sloan Research Paper 4732-09.
13. Meier, P. 2011. Why Bounded Crowdsourcing is Important for Crisis Mapping and Beyond. Available at http://irevolution.net/2011/12/07/why-bounded-crowdsourcing/
14. Poblet, M., Garcia-Cuesta, E., Casanovas, P. 2104. Crowdsourcing Tools for Disaster Management: A Review of Platforms and Methods. In *Lecture Notes in Artificial Intelligence*, Springer Verlag, Berlin, Heidelberg (forthcoming)
15. Poblet, M.; Leshinsky, R., Zeleznikow, J. 2012. Digital neighbours : Even Good Samaritan crisis mappers need strategies for legal liability, Planning News, 38, 11, Dec 2012: 20-21.
16. Quinn, J. A., Bederson, B. B. 2011. Human Computation: A Survey and Taxonomy of a Growing Field", CHI Conference, May 7-12, 2011 Vancouver, BC, Canada: 1403–1412. Available at http://alexquinn.org/papers/Human%20Computation,%20A%20Survey%20and%20Taxonomy%20of%20a%20Growing%20Field%20(CHI%202011).pdf
17. Saxton, G.D., Onook, O., Kishore, R. 2013. Rules of Crowdsourcing: Models, Issues, and Systems of Control Information Systems Management, Vol. 30(1): 2-20.
18. Schenk, E., Guittard, C. 2011. Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics* 1/2011 (n°7). Available at www.cairn.info/revue-journal-of-innovation-economics-2011-1-page-93.htm
19. Zhao, Y., Zhu, Q. 2012. Evaluation on crowdsourcing research: Current status and future direction. *Information Systems Frontiers.* Published online: 11 April 2012.