Learning from information crises: Exploring aggregated trustworthiness in big data production

Karin Hansson and Love Ekenberg
Stockholm University and the International Institute of Applied Systems Analysis
khansson@dsv.su.se, ekenberg@iiasa.ac.at

Abstract. In a crisis situation when traditional venues for information dissemination aren't reliable and information is needed immediately "aggregated trustworthiness", data verification through network evaluation and social validation, becomes an important alternative. However, the risk with evaluating credibility through trust and network reputation is that the perspective can get biased. In these socially distributed information systems there is therefore of particularly high importance to understand how data is socially produced and by whom.

The purpose with the research project presented in this position paper is to explore how patterns of bias in information production online can become more transparent by including tools that analyze and visualize aggregated trustworthiness.

The research project consists of two interconnected parts. We will first look into a recent crisis situation, the case Red Hook after Hurricane Sandy, to see how the dissemination of information took place in the recovery work, focusing on questions of credibility and trust. Thereafter, this case study will inform the design of two collaborative tools where we investigate how social validation processes can be made more transparent.

Social media in crisis management

Decision-making in self-organized systems such as social media can be understood as autonomous systems that goes beyond the centralized power of the nation-state, and where the network is the organizational principle. In this dynamic information production, decision-making takes place in the

collaborative, decentralized network of peers. Such tools are interesting when more traditional and central information channels have collapsed or are not perceived as trustworthy. Without functional public institutions providing reliable information, we seem to utilize other sources that we perceive as trustworthy, like family, friends and extended networks as well as new mechanisms, such as data verification through network evaluation. These issues are particularly during extreme events and crisis situations (Palen et al., 2010). It is against this, interesting to note that social media actually has, at many occasions, been demonstrated to be useful during crises, and in disaster contexts been useful for compensating for organizational as well as communication weaknesses (Palen & Liu, 2007). Social media has furthermore widely been adopted in voluntary organizations as a means to create civic engagement and organize collective action (Obar, Zube, & Lampe, 2012; Starbird & Palen, 2011), while authorities have been more cautious in their adoption of social media. The hurricane Sandy had a significant impact in this respect and led to a change in many officials' attitudes in New York and fostered an awareness of and interest in using social media to interact with the public (see, for example, McKay 2014; Sullivan and Uccellini 2012), but even though there is some interest in the possibilities of social media in crisis management there is a significant lack of structured approaches as well as tool support when it comes to important social and deliberative aspects of communication practices, such as coordination and bridge building (Voida, Harmon, & Al-ani, 2012), or to motivate long-term commitment to the civic sphere (Starbird & Palen, 2013). There are furthermore still severe problems due to a lack of trust and accountability (Antoniou & Ciaramicoli, 2013) and the highly related question about representation in social media practices is seldom touched upon at all.

A large issue here is that online networked publics quickly can become 'echo chambers' where people's opinions are reinforced by like-minded (Sunstein, 2001), and where this homogeneity is reinforced by information 'filter-bubbles' that mainly present information based on earlier searches (Pariser, 2011). To address these tendencies, new technologies have been researched and developed in different areas to promote deliberative publics where different opinions can meet. However these technologies are most often based on a simplified deliberative democratic ideal, where deep conflicts and dominating discourses are ignored. Efficient technology design to support equal representation and analysis of representativeness is lacking (Hansson, 2015).

Hurricane Sandy research

When looking at what has been published in connection with Hurricane Sandy there are two different ways of framing the crisis. The first is focused on big data to improve insights as well as providing more accurate future prognoses to enable

a better distribution of city resources when needed. See e.g.(Gupta, et al. 2013; Munro et al., 2013; Preis et al. , 2013). I this discourse, the crisis was an information problem solvable through greater transparency and public innovation. There is a great deal of research dealing with the quality of catastrophic data under various conditions and its trustworthiness from a technical perspective, such as spam detection, compromised accounts, malware and phishing attacks, but, the question of representativeness and information quality of the data is not handled systematically in any way.

The second way of framing the crisis focuses on the specific human values the crisis revealed regarding the community and the dependencies between different groups of people. Al-akkad et al. (2013) shows e.g. how people made creative use of the remains of the technological landscape. White, Palen, & Anderson (2014) point to the advanced collaborative work by the crowd using social media. News media also discussed the differences and inequalities the crisis revealed, which meant that it affected different people to different degrees depending of various socio-economical factors (Cher, 2012; Rohde, 2013). Thus, the presumable rational systems were not sufficient for dealing with the actual crisis and instead it turned out to be the fundamental need to care for each other and support their group that allowed people to manage catastrophic situations such as Hurricane Sandy. The normal communication technology had collapsed and people used available means for maintaining and establishing relations (Al-akkad et al., 2013). This way of framing the crisis defines infrastructures as relations and dependencies, and communication technologies as something used to strengthen mutual relations between people, actors and artifacts.

Project purpose and aims

In this project, we will focus on the social production of information, by exploring the qualitative base for data in the particular moment when the data is selected, collected and verified by situated subjects in networked publics built on dependencies. By developing tools that describe how the aggregation of trust takes place in these networked publics, biases might become more transparent, and thus enable a better understanding of the information reliability.

We will implement and evaluate this functionality in two tools:

* A crowdsourcing platform for evaluating and sharing of large amounts of information from informants during humanitarian crises, while investigating how social validation processes can be made clear in several different types of data feeds, and thus contribute to the evaluation of information reliability.
* A decision-support tool for online collaboration, while investigating how internal social validation processes can be clarified, and thus contribute to a better understanding of how groups of participants are represented in these dynamic decision processes.

References

Al-akkad, A., Ramirez, L., Denef, S., Boden, A., Wood, L., Büscher, M., & Zimmermann, A. (2013). "Reconstructing normality": The use of infrastructure leftovers in crisis situations as inspiration for the design of resilient technology. In *Proceedings of the 25th Australian Computer-Human Interaction Conference on Augmentation, Application, Innovation, Collaboration* (pp. 457–466). New York: ACM Press. doi:10.1145/2541016.2541051

Antoniou, N., & Ciaramicoli, M. (2013). Social media in the disaste cycle usefule tools or mass distraction? In *International Astronautical Congress*. Beijing.

Cher, L. (2012). Hurricane Sandy spotlights Brooklyn inequality | GlobalPost. *Global Post*. Retrieved February 22, 2014, from http://www.globalpost.com/dispatches/globalpost-blogs/groundtruth/hurricane-sandy-spotlights-inequality-red-hook

Gupta, A., Lamba, H., Kumaraguru, P., & Joshi, A. (2013). Faking Sandy: characterizing and identifying fake images on Twitter during Hurricane Sandy. In *CSCW 2013 Proceedings* (pp. 729–736). International World Wide Web Conferences Steering Committee. Retrieved from http://dl.acm.org/citation.cfm?id=2487788.2488033

Hansson, K. (2015). *Accommodating differences: Power, belonging, and representation online*. Stockholm University.

McKay, J. (2014). How Sandy changed social media strategies in New York City. *Emergency Management*. Retrieved Feb. 21, 2014, from http://www.emergencymgmt.com/disaster/Sandy-Social-Media-Strategies-New-York-City.html

Morstatter, F., Kumar, S., Liu, H., & Maciejewski, R. (2013). Understanding Twitter data with TweetXplorer. In *The 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (p. 1482). New York: ACM Press.

Munro, R., Erle, S., & Schnoebelen, T. (2013). Quality analysis after action report for the crowdsourced aerial imagery assessment following hurricane Sandy. In T. Comes, F. Fiedrich, S. Fortier, J. Geldermann, & L. Yang (Eds.), *Proceedings of the 10th International ISCRAM Conference* (pp. 10–11). Baden-Baden.

Obar, J. A., Zube, P., & Lampe, C. (2012). Advocacy 2.0: An analysis of how advocacy groups in the United States perceive and use social media as tools for facilitating civic engagement and collective action. *Journal of Information Policy*, *2*, 1–25.

Palen, L., Anderson, K. M., Mark, G., Martin, J., Sicker, D., Palmer, M., & Grunwald, D. (2010). A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters. In *Proceedings of ACM-BCS Visions of Computer Science 2010 1* (pp. 1–12). British Informatics Society.

Palen, L., & Liu, S. B. (2007). Citizen communications in crisis : anticipating a future of ICT-supported public participation. In *CHI 2007 Proceedings* (pp. 727–736). San Jose: ACM.

Pariser, E. (2011). *The filter bubble : what the Internet is hiding from you*. London: Viking.

Preis, T., Moat, H. S., Bishop, S. R., Treleaven, P., & Stanley, H. E. (2013). Quantifying the digital traces of Hurricane Sandy on Flickr. *Scientific Reports*, *3*, 3141. doi:10.1038/srep03141

Rohde, D. (2013). A year after Sandy, New York's inequality grows | David Rohde. *Reuters*.

Starbird, K., & Palen, L. (2011). "Voluntweeters": Self-organizing by digital volunteers in times of crisis. In *CHI '11* (pp. 1071–1080). Vancouver.

Starbird, K., & Palen, L. (2013). Working & sustaining the virtual " disaster desk ." In *CSCW '14: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 491–502). ACM.

Sullivan, K. D., & Uccellini, L. W. (2012). *Hurricane/post-tropical cyclone Sandy, Oct. 22–29, 2012*.

Sunstein, C. R. (2001). *Republic.com*. Princeton, N.J.: Princeton University Press.

Voida, A., Harmon, E., & Al-ani, B. (2012). Bridging between organizations and the public: Volunteer coordinators' uneasy relationship with social computing. In *CHI'12*. Austin.

White, J. I., Palen, L., & Anderson, K. M. (2014). Digital mobilization in disaster response: The work & self-organization of on-line pet advocates in response to hurricane Sandy. In *CSCW '14: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 866–876). Baltimore.