Linking and Visualizing
Social Media Data about Crises

Tomi Kauppinen

Department of Media Technology
Aalto University School of Science, Finland
tomi.kauppinen@aalto.fi

Abstract. Making sense out of data is a crucial process in decision making. This is especially so when it is for saving lives in crises and disasters like earthquakes, floodings or forest fires. Social media is a prominent source for human observations about disasters. Linking these crowdsourced observations together with SMS messages sent to aid phone numbers carry potential information about not only buildings collapsed or roads blocked, but also about people in need for medical services, food and water. This calls for linking the different pieces of information together by making temporal, spatial and thematic references explicit and groupable. We argue that this supports information usability and would thus create grounds for decision making. Examples of social media data and its handling about the earthquake in Haiti on January 12, 2010 serve in this to illustrate both the research problems and solutions.

1   From data to sensemaking

Crowdsourcing of Linked Open Data from social media about earthquakes and other disasters supports information filtering and grouping tasks [2]. These observations about the environment and humanitarian needs together with information about the three W's (3W)—*who* (which organizations) are doing *what* and *where*—together create a promising setting for improving timely decision making [1].

However, the essential challenge is that observational data coming from various sources (social media, SMS messages) are *raw* data, and thus need enrichment to be considered as usable information. For instance, Natural language processing (NLP) techniques are needed to reveal the underlying categories in messages and to encode them accordingly. An alternative is to rely on crowdsourcing efforts to provide the categorisation of messages—Ushahidi Platform[1] is an example of this approach. For the use in descriptions of the web of data, the category terms are best described as vocabularies.

Management of a Crisis (MOAC) vocabulary[2] does this by allowing to refer to category terms and properties via Uniform Resource Identifiers (URIs). Important here is the richness of sources for validating the genericness: the MOAC

---

[1] http://www.ushahidi.com/product/ushahidi/
[2] http://observedchange.com/moac/ns/

terms originate from various state-of-the-art sources including the Inter Agency Standing Committee (IASC), Emergency Shelter Cluster in Haiti, UNOCHA 3W Who What Where Contact Database and the Ushahidi platform.

Haiti earthquake being the first disaster where large scale crowdsourced information was gathered [3] serves as a case for understanding challenges one has about data. For this Linked Haiti is a dataset available and accessible online[3] for exploring and understanding the Haiti Earthquake in 2010. For instance, Figure 1 illustrates the power of grouping social media messages by categories and visualizing the results as interactive bubbles[4].
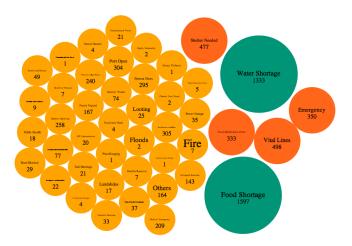

Fig. 1. Social Media Data about Haiti visualized as bubbles representing message categories. Sizes of bubbles indicate the amount of messages, accordingly.

References

1. Minu Limbu. Integration of crowdsourced information with traditional crisis and disaster management information using Linked Data. Master's thesis, University of Muenster, Germany, 2012.
2. Jens Ortmann, Minu Limbu, Dong Wang, and Tomi Kauppinen. Crowdsourcing Linked Open Data for disaster management. In Rolf Grütter, Dave Kolas, Manolis Koubarakis, and Dieter Pfoser, editors, *Proceedings of the Terra Cognita Workshop on Foundations, Technologies and Applications of the Geospatial Web, In conjunction with the International Semantic Web Conference (ISWC2011)*, volume 798, pages 11–22, Bonn, Germany, October 2011. CEUR Workshop Proceedings.
3. Matthew Zook, Mark Graham, Taylor Shelton, and Sean Gorman. Volunteered Geographic Information and Crowdsourcing Disaster Relief: A Case Study of the Haitian Earthquake. *World Medical & Health Policy*, 2(2):7–33, 2010.

---

[3] http://linkedscience.org/data/linked-haiti/

[4] This technique is realized by the help of the D3.js library