The Use of Social Media within the Global Disaster Alert and Coordination System (GDACS)

Beate Stollberg
Joint Research Centre of the European Commission,
Institute for the Protection and Security of the Citizen,
Via E. Fermi, 21027 Ispra, Italy
beate.stollberg@jrc.ec.europa.eu

Tom de Groeve
Joint Research Centre of the European Commission,
Institute for the Protection and Security of the Citizen,
Via E. Fermi, 21027 Ispra, Italy
tom.de-groeve@jrc.ec.europa.eu

ABSTRACT
The Global Disaster Alert and Coordination System (GDACS) collects near real-time hazard information to provide global multi-hazard disaster alerting for earthquakes, tsunamis, tropical cyclones, floods and volcanoes. GDACS alerts are based on calculations from physical disaster parameters and used by emergency responders. In 2011, the Joint Research Centre (JRC) of the European Commission started exploring if and how social media could be an additional valuable data source for international disaster response. The question is if awareness of the situation after a disaster could be improved by the use of social media tools and data. In order to explore this, JRC developed a Twitter account and Facebook page for the dissemination of GDACS alerts, a Twitter parser for the monitoring of information and a mobile application for information exchange. This paper presents the Twitter parser and the intermediate results of the data analysis which shows that the parsing of Twitter feeds (so-called tweets) can provide important information about side effects of disasters, on the perceived impact of a hazard and on the reaction of the affected population. The most important result is that impact information on collapsed buildings were detected through tweets within the first half an hour after an earthquake occurred and before any mass media reported the collapse.

Categories and Subject Descriptors
H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing – *Linguistic processing.*

General Terms
Experimentation, Security, Human Factors.

Keywords
Social Media, Twitter, Disaster Management, Emergency Response, Impact Analysis.

WWW 2012 Companion, April 16–20, 2012, Lyon, France.
ACM 978-1-4503-1230-1/12/04.

1. INTRODUCTION
The Global Disaster Alert and Coordination System (GDACS) provides global multi-hazard disaster monitoring and alerting for earthquakes, tsunamis, tropical cyclones, floods and volcanoes. It is a joint initiative of the United Nations Office for the Coordination of Humanitarian Affairs (UN OCHA) and the European Commission (EC) that serves to consolidate and improve the dissemination of disaster-related information, in order to improve the coordination of international relief efforts. Through partnerships with scientific organisations and other hazard monitoring institutions, GDACS collects near real-time hazard information, which is combined in Geographic Information System (GIS) models with demographic and socio-economic data. Disaster alerts and impact estimations for natural disasters are automatically disseminated minutes after disaster events. Estimations are calculated from disaster parameters and various databases (population density, vulnerability, etc.). This service is provided by the European Commission Joint Research Centre (JRC) and is improved continuously.

In 2011 JRC started to explore if social media could be an additional valuable data source for decision makers who are using GDACS. Social media is defined as a group of Internet-based applications that build on the ideological and technological foundations of the Web 2.0, which allows the creation and exchange of user-generated content (UGC) [1]. This means the content of websites is not only published by the administrators of an application anymore but also by the end-users which therefore become information producers and consumers at the same time. The maybe most popular example for this trend is the free encyclopedia Wikipedia but also sites like YouTube for videos, Flickr for photos, and in general blogs, podcasts and so on. For the Geographic Information (GI) field this trend must be seen in connection with the development and spread of GPS devices. These became smaller and cheaper within the last years and a nowadays even integrated in many phones. Already in 2007 Goodchild [2] introduced the term Volunteered Geographic Information (VGI) and pointed out how citizens will more and more act as sensors by sharing their observations about their adjacencies.

JRC started to explore these current technology and internet trends by following three different approaches using social media: information dissemination, information monitoring and information exchange. These different communication schemata are described in chapter 2. Chapter 3 presents the developed application for parsing Twitter messages (so-called tweets) and the results of the data analysis. A summary and an outlook to future work are given in chapter 4.

2. COMMUNICATION SCHEMATA

The use of social media relevant for GDACS can be categorized into three fields, based on the communication schema: information dissemination, information monitoring and information exchange.

2.1 Information Dissemination

Social media can provide alternative if not more accurate and/or timely information compared to mass media and authorities for the affected population after a disaster occurred [3]. GDACS provides such important information in crisis situation. Until now this was done through "traditional media" like the GDACS website (http://gdacs.org) and the sending of alerts via SMS, Fax and Email. Social media can be used to spread this kind of information more widely and more targeted to interested communities using the social media model. In this case the communication is taking place in one direction, meaning that GDACS disseminates information to the users (one-to-many-communication). In order to do this, a Twitter account and a Facebook page were set up.

2.2 Information Monitoring

For emergency managers, social media have proven to add value in the detection of disasters and situation awareness (e.g. [4] for floods and [5] for forest fires). More importantly, social media are a source of information on the perceived impact of a hazard, on the reaction of the affected population and offers alternative views on the situation from various actors [6]. In this case the communication is also taking place in one direction, namely GDACS monitors user information (many-to-one-communication). For this reason a Twitter parser was developed which is the essential part of this paper and presented in chapter 3.

2.3 Information exchange

GDACS aims to exchange information with the users and to increase the collaboration between user groups. In this case the communication is taking place in both directions (many-to-many-communication). In May 2011 an Android App under the GEO-PICTURES project by AnsuR and UNITAR/UNOSAT[1] for geo-tagged photos was released to the disaster management and early responders' community. The developed application automatically geo-locates photos and sends them off to a web-server. On the server, the photo location is indicated on a map, and the photo can be further assessed and shared with individuals or groups.

Also in 2011, an exploratory research project was started at JRC which revolves around the development of an iPhone App to provide users with real-time information about disasters (information dissemination) and give them the possibility to send information in the form of a geo-located image and/or text back (information exchange) (Figure 1). The goal of the project is to extract and feed back useful information from the reports shared by the community for improving situation awareness and providing ground truth for rapid satellite-based mapping. The public launch of the App is planned for the beginning of 2012. Further work will address the processing of information for extracting added value information, comparable to the Twitter data analysis described in chapter 3.

[1] GEO-PICTURES crowd-sourcing application available to improve field assessment. Retrieved January 23, 2011, from http://www.unitar.org/unosat/node/22/1429.

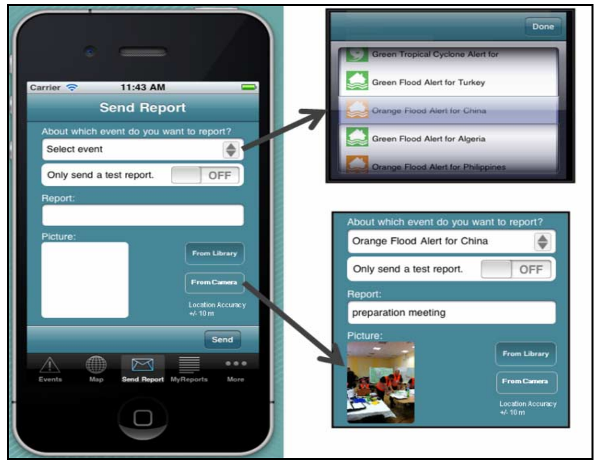

Figure 1 - Sending Reports through the Mobile Application

3. TWITTER PARSER

In October 2011, JRC set up a Twitter parser which continuously requests tweets containing the English keyword "earthquake" and stores them in a database. More than 1.2 million tweets were stored and analyzed for the period from the end of October until the middle of January 2012.

The first step was to analyze the overall number of tweets and how this number is varying day by day. Figure 2 shows a graph with the total number of tweets containing the word "earthquake" between 28.10.2011 and 17.01.2012.

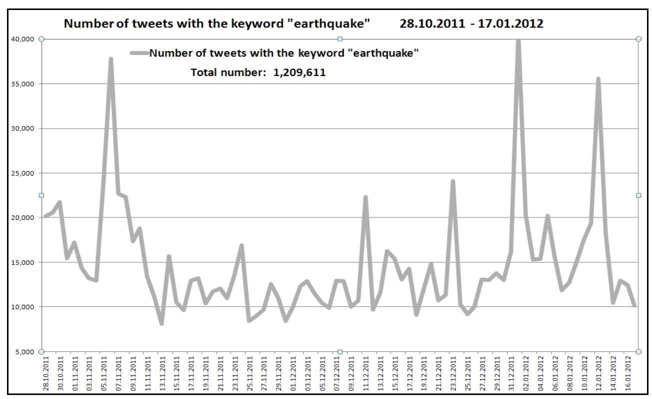

Figure 2 - Total Number of Tweets with the Keyword "earthquake" between 28.10.2011 – 17.01.2012

The graph shows that the number of tweets per day is between 8,000 and 25,000, apart from three big peaks: there were around 37,000 tweets at the 6th of November 2011, around 40,000 tweets at the 1st of January 2012 and around 35,000 tweets at the 12th of January 2012. Analyzing the tweets for these three specific days showed that the words which occurred most in these earthquake tweets were "Oklahoma" (6.11.11), "Japan" (1.1.12) and "Haiti" (12.1.12). Comparing these facts to the GDACS earthquake archive showed that there was a strong earthquake with magnitude 5.6 in Oklahoma on the 6.11.11 and one with magnitude 6.8 in the Japan region Izu Islands on the 1.1.12. There was no strong earthquake on the 12.1.12 in the Haiti region, but the 12th of January is the anniversary of the catastrophic earthquake in Haiti in 2010. In conclusion, these three peaks show that the (past) occurrence of strong earthquake events is very well reflected in the analyzed data.

The next graph (Figure 3) presents the number of tweets containing in addition to the word "earthquake" also the word "collapse" in the same period.

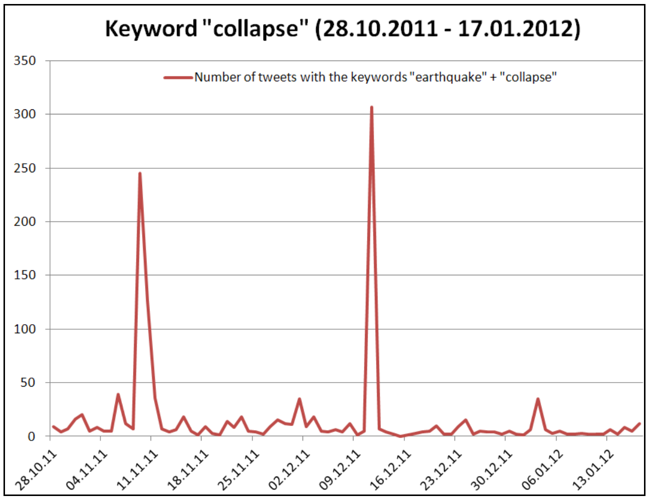

Figure 3 – Number of Tweets with the Keywords "earthquake" and "collapse"

This graph shows a peak on the 9th of November 2011 and the 11th of December 2011. Analyzing media reports for these days, it can be found out that on the 9th of November an earthquake with magnitude 5.6, an aftershock of the larger Van earthquake, caused a six-floor hotel to **collapse** in the Eastern Turkish province Van. An unknown number of people was trapped inside. Furthermore, on the 11th of December a strong earthquake with magnitude 6.5 (causing an orange earthquake alert) hit Mexico. An 11-year-old boy died in a roof **collapse** caused by the earthquake. These two events were (re)tweeted continuously and led to these peaks in the number of tweets for these specific days.

Figure 4 shows a graph with the number of tweets containing in addition to the word "earthquake" also the word "nuclear".

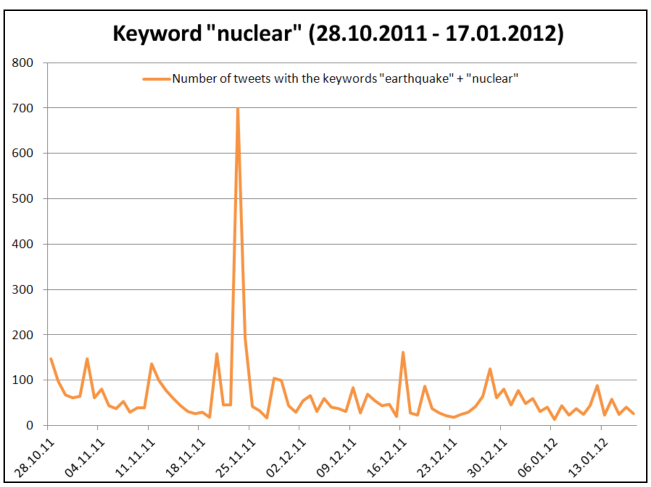

Figure 4 – Number of Tweets with the Keywords "earthquake" and "nuclear"

This word occurs normally around 20 – 150 times per day in the earthquake tweets, but there is one peak on the 23rd of November 2011. The GDACS earthquake archive shows that on this day a strong earthquake with magnitude 6.1 hit Japan, in the same area as the catastrophic earthquake in March 2011 which caused the tsunami destroying the **nuclear** power plant in Fukushima. This means that the peak does not indicate an actual nuclear accident but people are obviously afraid that an incident in the past could happen again.

Figure 5 shows a graph with the number of tweets containing in addition to the word "earthquake" also the word "landslide". There is one peak on the 5th of January 2012. This day an earthquake caused a **landslide** in the Compostela Valley on the Philippines. Dozens of people were killed this day. The other peak on the 16th of November 2011 does not lead back to a real landslide event. This day an American poet and musician tweeted the following sentence: *"...by a* ***landslide****, a mudslide, a tsunami, an earthquake, a hurricane, a tornado, all metaphors of vibrant emotion and epiphany."* This tweet was re-tweeted by many of his followers and caused the peak in the earthquake data.

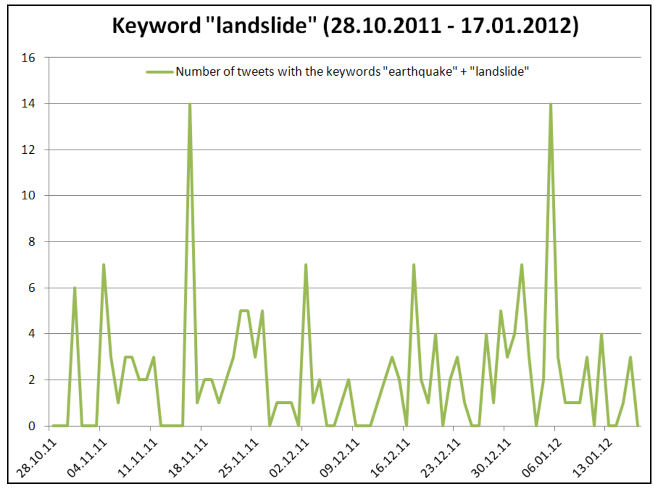

Figure 5 – Number of Tweets with the Keywords "earthquake" and "landslide"

Normalizing the number of tweets for specific impact keywords can show which ones actually occurred after an earthquake. which We call this method the creation of an "Impact Index" Figure 6 shows the normalized values for the keywords "aftershock", "collapsed", "earthquake", "hospital", "landslide", "looting", "nuclear", "prison" and "tsunami" for the strong aftershock on the 9th of November in Turkey described earlier. It shows a clear anomaly for the keyword "collapsed".

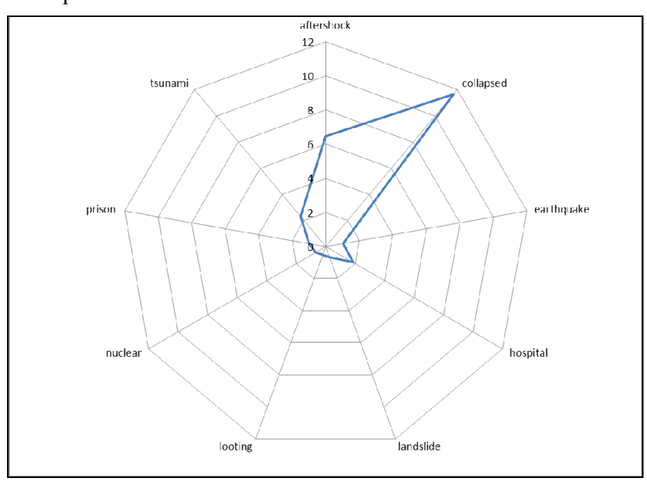

Figure 6 - Normalized Number of Tweets around the 9th of November 2011 Aftershock in Turkey

A more detailed analysis of tweets just after an earthquake shows that unique information on consequences and impact is available, before mass media reports are issued. The strong aftershock in Turkey on the 9th of November 2011 occurred at 20.23 MEZ. The following tweets (in English) were sent in the immediate aftermath:

- *"Earthquake in Van. 5.6. İt is unbelievable... Hotel collapsed. No more news."* (20:41 MEZ)
- *"Urgent:An earthquake measuring 5.6 on the Richter scale hit the province and that the Turkish and reported the collapse of several buildings"* (20:48 MEZ)
- *"Biggest afterschock earthquake of magnitude 5.6 at the province of Van, T Turkey. There are buildings collapsed. So sad."* (20:53 MEZ)

These three tweets were sent before the first Associated Press (AP) report was tweeted at 20:57 MEZ. After this, the "official" report was re-tweeted again and again, leading to the high numbers shown in Figure 7. This means that traditional media reports were available half an hour after the event and the interesting time frame is the first half an hour after an earthquake occurred. This is a very narrow window, where messages can only come from people affected by the earthquake. The tweets within this window can provide important information for emergency responders since no other source is available at this time.

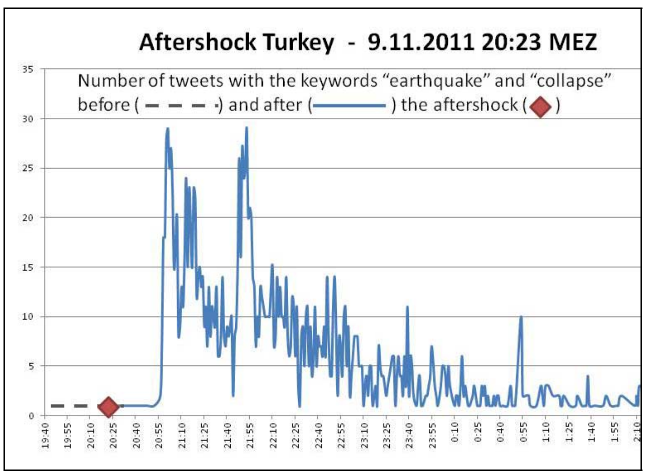

Figure 7 - Detailed Number of Tweets with the Keyword "collapse" around the strong Aftershock in Turkey on 9.11.11

4. CONCLUSION AND FUTURE WORK

JRC is following three different approaches to use social media within GDACS for providing added-value information to emergency responders: information dissemination, information monitoring and information exchange. A Twitter account was set up and GDACS alerts are now not only sent via traditional media like SMS, Fax and Email, but also through Twitter. Furthermore a Facebook page for data dissemination and exchange was developed and mobile applications for exchanging information with the users were implemented.

The essential part of this paper was the presentation of a developed Twitter parser for information monitoring and the analysis of the result data. The volume of Twitter messages after an earthquake indicates the interest of the Twitter community in the event. This shows that the occurrence of earthquakes is reflected in the number of tweets in general and it would be possible to detect a disaster by observing the number of tweets containing specific keywords. But since such detection can already be (better) done by physical sensors, this fact is only of secondary importance, we call it the calculation of an "Interest index". As demonstrated before, this can always be for a real event, or for an anniversary of an event. This means tweets also reflect what people are afraid of (like another nuclear accident in Japan) and which events they will always remember and have in mind (like the Haiti earthquake in 2010). For any single case it has to be analyzed and judged if people are tweeting about real occurring events or remind on past ones. At the very least, the absence of interest can be an indication that a strong earthquake did not cause a humanitarian disaster.

More interesting is that the impact of these earthquakes can be analyzed and possible side effects detected; we call this "Impact index". So far we could show that collapsed buildings and landslides "appeared" within the tweets as anomalies in Twitter messages. Furthermore, the very narrow time window before mass media reports are available and where messages can only come from people affected by a disaster, can provide important information for emergency responders.

The next step is to extend the Twitter parser by more disaster keywords (e.g. "flood", "tsunami", "storm", "hurricane" etc.) and by more languages. Afterwards the analysis will be automatically included into the GDACS website (http://gdacs.org) as an additional valuable data source for the international disaster response community.

5. REFERENCES

[1] Kaplan, A. M. and Haenlein, M. 2010. *Users of the world, unite! The challenges and opportunities of Social Media*. Business Horizons 53(1): 59–68.

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

[3] Sutton J., Palen, L. and Shklovski, I. 2008. *Backchannels on the Front Lines: Emergent Uses of Social Media in the 2007 Southern California Wildfires*. Proceedings of the 5th International ISCRAM Conference – Washington, DC, USA, May 2008.

[4] De Longueville, B., Luraschi, G., Smits, P., Peedell, S. and De Groeve, T. 2010. *Citizens as Sensors for Natural Hazards: a VGI Integration Workflow*. GEOMATICA Vol. 64, No. 1, 2010, 41-59.

[5] De Longueville B., Smith, R. and Luraschi, G. 2009. *"OMG, from here, I can see the flames!": a use case of mining location based social networks to acquire spatio-temporal data on forest fires*. Proceedings of the 2009 International Workshop on Location Based Social Networks (LBSN '09). ACM, New York, NY, USA, 73-80.

[6] Dandoulaki M. and Halkia, M. 2010. *Social Media (Web 2.0) and Crisis Information: Case Study Gaza 2008-2009*. In Advanced ICTs for Disaster Management and Threat Detection: Collaborative and Distributed Frameworks. Eds. Eleana Asimakopoulou and Nik Bessis, 143-163.