Leveraging Geospatially-Oriented Social Media Communications in Disaster Response

Susannah McClendon
GeoVISTA Center
Department of Geography
The Pennsylvania State University
susantbm@gmail.com

Anthony C. Robinson
GeoVISTA Center
Department of Geography
The Pennsylvania State University
arobinson@psu.edu

ABSTRACT

Geospatially-oriented social media communications have emerged as a common information resource to support crisis management. Our research compares the capabilities of two popular systems used to collect and visualize such information - Project Epic's Tweak the Tweet (TtT) and Ushahidi. Our research uses geospatially-oriented social media gathered by both projects during recent disasters to compare and contrast the frequency, content, and location components of contributed information to both systems. We compare how data was gathered and filtered, how spatial information was extracted and mapped, and the mechanisms by which the resulting synthesized information was shared with response and recovery organizations. In addition, we categorize the degree to which each platform in each disaster led to actions by first responders and emergency managers. Based on the results of our comparisons we identify key design considerations for future social media mapping tools to support crisis management.

Keywords

Geographic Information, Social Media, Crisis Management, Mashups.

INTRODUCTION

Crowd-sourced information has rapidly become an essential source of data in disaster response. Since the first well documented efforts of citizen journalists on September 11th, 2001 and the use of internet blogs to collect information after the 2004 East Indian Ocean Tsunami, recent emergency response efforts have included mapping SMS messages after the Haiti earthquake in 2010. The Haiti earthquake represented a paradigm shift in the use of social media for disaster response, as multiple web-based platforms emerged to collect, refine, and disseminate crisis-related social media. The use of social media to gain real time information on the ground in a disaster has been driven by the rapid speed at which information can be distributed, the cross-platform accessibility of information, and the ubiquity of social media worldwide (Vieweg, et al., 2010). The utility of this information has been enhanced by the creation of crisis maps based on location data extracted from social media communications (Liu and Palen, 2010, MacEachren, et al., 2011).

In 2011 the American Red Cross conducted a survey that showed that 33% of citizens have used social media sites, including Facebook, Twitter, Flickr and SMS text messages/alerts to gain information about an emergency (American Red Cross, 2011). About half the respondents said they would contribute information during an emergency using social media channels. Statistics from the International Telecommunication Union reveal that in 2009 there were 4.6 billion mobile phone subscribers world-wide and 1.5 billion subscribers used mobile devices to access the internet (International Telecommunication Union, 2011). We can reasonably expect the use of social media in disaster response to increase in the future.

Research on the increasing use of social media in disaster response has emerged as a new focus in the field of crisis informatics (Anderson and Schram, 2011). Extracting, categorizing, visualizing, and evaluating such information presents serious research challenges, including the problem of managing and extracting meaningful information from the large volume of contributions, applying the information to decision support workflows, and the development of formal information sharing protocols (Harvard Humanitarian Initiative, 2011). Mapping crowd sourced information in disaster response gained wide-scale media attention after the successful deployment of the Ushahidi Crisis Map during the 2010 Haiti earthquake (Starbird, 2011). There are a number of specific challenges involved in mapping social media communications, including the extraction of accurate location information, and the application of useful and usable cartographic representations to visually support situational awareness in crises. Research on the integration of Geographic Information Systems (GIS) and crowdsourced information from social media has focused more on the challenges of extracting action items and location information from social media feeds (MacEachren, et al., 2011) and less on the utility of the extracted information and the effectiveness of associated crisis maps to support emergency response.

Our research examines two applications that have leveraged geospatially-oriented social media during recent disasters; the Tweak the Tweet (TtT) project from Project Epic (Starbird and Palen, 2011) and Ushahidi (Okolloh, 2009), both of which have been used to create crisis maps of content collected from social media sources during recent disasters. For each application we examine collected data, information products, and evidence of subsequent response actions for two recent disasters; the 2011 Joplin tornado and 2010 Fourmile Canyon fire for Tweak the Tweet, and the 2010 Haiti earthquake and 2010 Gulf Oil Spill for Ushahidi. Other efforts have used crowdsourced information during recent disasters, including the open source Sahana platform (Currion, et al., 2007) and the collective effort of the Crisis Mappers Network (crisismappers.net). However, Sahana utilized data feeds directly from Ushahidi and TtT, and the Crisis Mappers Network is focused on connecting and empowering crisis collaborators, and does not offer their own specialized technology platform.

We begin with background information, including an overview of geospatially-oriented social media, followed by a brief history of TtT and Ushahidi. Next, we evaluate how each organization collected, processed and geo-located these social media communications and compare and contrast the cartographic representations and reporting capabilities of the resulting crisis maps. We then examine the effectiveness of each application by identifying examples of actionable items used by military, government and non-government organizations that emerged from the use of these crisis maps. Finally, we conclude with key design considerations for future efforts to leverage geospatially-oriented social media in crisis informatics.

GEOSPATIALLY-ORIENTED SOCIAL MEDIA

Many social media sources, including Twitter, allow users to tag reports with coordinates to indicate their location on Earth. This information is easy to process and represent on a map, but does not necessarily represent an “actionable” location. A more substantive challenge is associated with making use of textual descriptions of place (placenames and less-specific geographic features), like those often included in an SMS text message. Placenames in text can be geocoded to assign location coordinates, but placenames are usually associated with irregular areas (for example, the New York City metro area) at least as often as one might ever associate them with a specific coordinate location on earth (the centroid of the legal boundary of New York City). Determining less-specific geographic features is also of critical importance when using geospatially-oriented social media.

Both platforms make use of social media collected from SMS and Twitter messages. SMS or “text messaging” is a short messaging service that allows for the storage and retrieval of short 160 character messages across global cellular telephone networks. Location information is not automatically attachable to SMS data, and must be inferred from the message itself. Twitter is a microblogging service that allows users to post messages up to 140 characters called Tweets via mobile phones or web accessible devices. Twitter users follow other users to see their tweets in a Twitter feed. The Twitter user community has developed linguistic markers to facilitate communication; including the @ symbol to address users (@username); the RT abbreviation to represent a retweet (RT @username); and # or hashtag to indicate keywords. Hashtags allow users to search Twitter feeds or to follow trends (Zappavigna, 2011). Location information can be added to a tweet using a phone’s GPS capabilities, and can also be inferred from user profiles and mentions of placenames in messages themselves.

PROJECT EPIC AND TWEAK THE TWEET

Project Epic is research effort at the University of Colorado that aims to improve methods of public information gathering and dissemination during emergency situations. The project’s mission is to couple computational methods with behavioral knowledge on how people develop information using social media in crisis situations (Palen, et al., 2010). One Project Epic research project, Tweak the Tweet, was first presented in 2009 as a simple set of standardized communication practices coupled with a technology platform for making sense of crisis Tweets (Starbird, 2011, Starbird and Palen, 2011). TtT asks users to tweet using a crisis specific micro-syntax designed to enable real-time processing of Tweets. TtT features a web-based tool for collecting and visualizing contributed information using the Twitter API to continually-update a database. The categorized information is displayed on a simple map mashup using the Google Maps API.

The TtT micro-syntax is based on primary or main hashtags that can be used in any crisis situation and are designed to indicate the "who, what, and where" of the Twitter message content. For example, #name or #contact can be used to indicate "who"; #need, #shelter, #road, #open, #damaged can be used to indicate "what"; and #loc can used indicate "where". These hashtags are used in conjunction with an event tag to organize the crisis. Event tags can be spontaneously generated during an event, like #joplin or #tornado, or prescribed by the TtT micro-syntax like #4MileFire. Used together, the primary and event hashtags format meaningful machine readable tweets.

TtT was first deployed in the aftermath of the Haiti earthquake in January 2010 with the goal of having responders and agencies on the ground use the syntax. The first deployment did not have any associated mapping functionality, and the micro-syntax was not widely adopted by first responders or the public. Despite that, volunteers from Crisis Commons, TtT and other organizations tweeted or retweeted almost 3000 unique tweets formatted with the TtT syntax (Starbird and Palen, 2011). Since 2009, TtT has added a mapping component to their system design and the application has been deployed for over twenty major crises.

USHAHIDI

Ushahidi began as a non-profit African technology company that was developed to map incidents of violence in Kenya following elections in 2008. Ushahidi's mission is to develop platforms for sharing crisis information and personal narratives (Okolloh, 2009) and has since grown to develop tools to facilitate the democratization of information in broader contexts. The open source software tools developed by Ushahidi automate the collection of incident reports using cellular phones, email, and the web and facilitate the mapping of report locations in an interactive map mashup along with descriptive data to contextualize events.

Ushahidi offers three core products: the Ushahidi Platform, the SwiftRiver Platform, and Crowdmap. The Ushahidi Platform combines interactive mapping with the ability to capture real-time data streams from mobile messaging services and Twitter, and also supports email and web forms. It also provides spatial and temporal views of collected data. The SwiftRiver Platform allows for the real-time filtering and verification of data from these multiple data streams, including the ability to automatically categorize information based on semantic analysis, provide analytics and insight into user relationships and data trends, facilitate information validation and qualification, and it offers an interactive dashboard for monitoring and reporting purposes. Crowdmap is a cloud-hosted solution designed to support rapid launches of both the Ushahidi and SwiftRiver platforms.

Since 2009, deployments of Ushahidi platforms have focused on election monitoring, reporting human rights violations, disease surveillance, wildlife tracking, and disaster response. Though there were several deployments of the Ushahidi platform prior to the 2010 Haiti earthquake, it was the Haiti crisis that brought Ushahidi international attention. Ushahidi adoption since the 2010 Haiti earthquake has seen significant growth.

DATA COLLECTION, MANAGEMENT, AND IDENTIFYING LOCATIONS

Both TtT and Ushahidi utilize technical and manual methods to collect, refine, and add meaning to data. The following sections describe how each platform is designed, how they manage data, and how they derive location information from collected social media reports.

Tweak the Tweet

After the initial launch of TtT during the 2010 Haiti earthquake, TtT refined its aims to promote *crowdfeeding* after analysis of results from the deployment in Haiti highlighted the difficulty of getting the crowd to adopt a micro-syntax for Twitter (Starbird and Palen, 2011). TtT promotes the monitoring of social media sites by volunteers (called voluntweeters) during a crisis and they disseminate information back into the crowd using the TtT micro-syntax. In addition, voluntweeters promote the use of the syntax through conversations with other responding organization volunteers and by posting instructions and links to TtT crisis maps on social media sites. For the events we researched, TtT prescribed a micro-syntax with event tags including #boulderfire, #boulder, #4MileFire, #joplin, and #tornado.

The TtT software platform utilizes the Twitter Streaming API to identify tweets based on the TtT micro-syntax and stores the tweets in a MySQL database, parsing the information into key-value pairs based on hashtags. The platform uses Google maps to map the tweet content after the location has been determined. At regular intervals a Ruby script parses the messages filtered by hashtags into a MySQL database and the script in turn updates a public Google spreadsheet. Because machine processing may miss meaningful data in the tweet, such as placenames and other locations, the TtT process uses a combination of automatic and manual processing by volunteers to populate data in the event spreadsheet. For this research we downloaded spreadsheets for the 2010 Boulder Fourmile Canyon fire and the 2011 Joplin tornado disasters. These TtT–hosted spreadsheets contain all the events collected for each disaster.

Ushahidi

In the hours following the 2010 Haiti earthquake, Ushahidi staff deployed the Ushahidi Haiti Crisis Map. Working in the United States, they gathered information from media reports and social media sources. Approximately 85% of Haitians had access to cellular telephones and the cellular telephone infrastructure, though damaged, was quickly repaired. Within days a SMS short code number was set up in collaboration with phone companies and U.S. State Department resources and advertised through local radio stations. The messages received via SMS were sent to an automated system set up to facilitate message translation and mapping of the data by volunteers (Heinzelman and Waters, 2010).

Shortly before the 2010 Gulf Oil Spill, students at Tulane University began development of a crisis map to document oil refinery accidents using the Ushahidi platform. On the day the class presented the GIS map, the Deepwater Horizon oil rig exploded in the Gulf of Mexico (Dosemagen, 2010). The Louisiana Bucket Brigade, an environmental organization, worked with Tulane students to launch the Oil Spill Crisis Map to give Gulf residents a chance to contribute information about threats to their community and ecosystem from the oil spill. Data for the map was submitted via SMS, Email, Twitter and web forms. Citizens were encouraged to make a reports based on health issues, wildlife sightings, and other notable impacts they may witness in the region.

The Ushahidi API (Ushahidi, 2011) supports data exchange in XML (Extensible Markup Language) and JSON (JavaScript Object Notation). Ushahidi software supports PHP scripting and is designed to work with MySQL and is usually run on an Apache web server. Ushahidi software can be configured to work with common SMS gateway providers to process and deliver SMS messages, and it can be configured to use the Twitter Streaming API to process Tweets. Data can be exported from MySQL via a PHP script to a Google spreadsheet. The Ushahidi map template is designed with a link to download the raw data in a Google spreadsheet, but because the Ushahidi platform is open source and can be modified by the organizations that deploy the software, not all organizations include the ability to download the data. For the 2010 Haiti earthquake and the 2010 Gulf Oil Spill, we were able to download spreadsheets with data covering six months after the initial incidents.

Streaming Data and Scalability Challenges

Collecting data from social media communications like Twitter and SMS is difficult due to the large datasets that can be generated in a short amount of time. A key challenge has emerged in automating the extraction of useful and actionable data from such sources. In fact, applying structure to content using tweets with a micro-syntax to enhance computational automation was part of the original intent behind the TtT project. Challenges associated with filtering, managing, analyzing and translating large volumes of social media communications are being addressed through ongoing development of The SwiftRiver platform by Ushahidi.

During the first deployment of TtT for the 2010 Haiti earthquake, the syntax was not widely adopted by citizens and first responders, but the syntax was picked up by people who spontaneously volunteer during a crisis (Starbird and Palen, 2011). TtT efforts spurred a network of volunteers that helped give structure to the social media communications that were transpiring on Twitter during both the 2011 Joplin tornado and 2010 Boulder Fourmile Canyon fire crises. These volunteers adopted the TtT syntax and translated information from multiple sources using the syntax before tweeting it out to their followers. These followers were diverse, including media outlets, the American Red Cross, FEMA, and other relief organizations. This type of volunteerism was promoted to direct Twitter communications so that automatic filtering of Tweets would be more effective.

During the 2010 Haiti earthquake, Ushahidi enlisted volunteers to assist with handling the large volume of data. SMS messages began to flow at a rate of 1,000 to 2,000 a day and were passed directly from the cellular telephone provider to an automated system, designed by Ushahidi developers for coordinating volunteers. Volunteers manually translated the messages from Haitian Creole and then filtered and determined locations (Meier and Munro, 2010). The system supported message translation with a lead time of less than ten minutes.

EXTRACTING LOCATION INFORMATION

One of the most challenging aspects of using social media data during a disaster is extracting unambiguous and accurate location information. Locations are essential for determining if a message is actionable (Munro, 2011). Location can be determined in several ways, including processing location references like a place name or street address in message content; explicit coordinates derived from geo-location services from cellular phones; and extraction of location information in a user's social media account profile (Bellucci, et al., 2010, Field and O'Brien, 2010). Table 1 illustrates examples of profile-derived location information shown in Tweets 5-8.

Tweet	Time	User	Tweet	Location
1	11.21 pm 23 May 2011	@sarahgracesitz	RT @shawncmatthews: Here is a great resource for donation centers #joplin #tornado #relief http://ow.ly/515zC	iPhone: 37.112511,-93.303925
2	11.18 pm 23 May 2011	@CajunTechie	If you want to donate clothing of all sizes to displaced residents in #Joplin #Missouri you can drop it by 702 Moffet #tornado	ÃœT: 36.874136,-94.873582
3	10.07 pm 23 May 2011	@pbdoetmee	Ronduit dramatisch fotowerk na de tornado hit-down in Joplin, Missouri, USA... http://bit.ly/jiw6Kg #indrukwekkend #joplin #tornado	51.953923,6.008155
4	11.35 pm 23 May 2011	@PeterKinder	Just confirmed I will be guest on @IngrahamAngle radio show Tuesday morn 5/24/11 9:30 CDT talking #Joplin #MO #tornado relief, update #pdk	iPhone: 0.000000,0.000000
5	10.48 pm 23 May 2011	@maryfranholm	RT @OzarksRedCross: #CBCO FB site says: CODE RED 4 blood donations 4 the #Joplin #tornado has been lifted. Thanks 2 so many of you donat ...	Lost in a good book
6	10.39 pm 23 May 2011	@Jeannie_Hartley	RT @OzarksRedCross: #RedCross update here: http://bit.ly/jZ3Ivp #Joplin #tornado	Universe
7	10.36 pm 23 May 2011	@wheelertweets	RT @Jeannie_Hartley: #Tornado #Joplin #mo @info4disasters @Redcross @kcredcross @1stAid4 @wheelertweets @jnicky63 @viequesbound @Lady1st ...	Tunis, Tunisia
8	11.02 pm 23 May 2011	@JoplinMoTornado	GOODNEWS: 7 people were rescued from the debris today! #joplin #tornado -G	Joplin, Mo
9	11.02 pm 23 May 2011	@DAOWENS44	RT @OzarksRedCross: #CBCO FB site says: CODE RED 4 blood donations 4 the #Joplin #tornado has been lifted. Thanks 2 so many of you donat ...	

Table 1. Tweet Examples from the 2011 Joplin Tornado.

Location information included within a Twitter or SMS message as a text reference (e.g. a user mentions a specific place by name) must be extracted and geocoded to obtain coordinate information. Latitude and Longitude coordinates can also be included with messages as geospatial metadata. The process of manually or computationally assigning such metadata is called geo-tagging. Twitter included the ability to geo-tag tweets in 2009 (Bellucci, et al., 2010). Because of privacy concerns, social media applications and cellular phones usually require users to opt-in to enable geo-tagging. The SMS protocol does not incorporate geospatial metadata and typical messages sent from cellular phones via SMS will not contain location information (Munro, 2011). However, GeoSMS (geosms.wordpress.com), a location-enabled SMS standard, can embed geospatial metadata into a URI (Uniform Resource Identifier). MacEachren, et al. (2011) notes that the proportion of users who enable geo-tagging is still small. However, geo-tagging alone is no guarantee the message content is meaningful. Of three geo-tagged examples in Table 1, two geo-tags are close to Joplin in Springfield, Missouri (1) and the nearby city of Miami, Oklahoma (2), while the location in Tweet 3 is in the Netherlands.

The Ushahidi platform does not contain a mechanism to automatically geocode implicit location information, but the SwiftRiver Platform does incorporate tools that use natural language processing and a gazetteer to return coordinate locations based on place names. The Ushahidi platform will extract geospatial metadata from social media feeds if it exists. For the Ushahidi 2010 Haiti earthquake map, the majority of information gathered came via SMS and not geo-tagged. Location information from SMS messages was translated by volunteers who used a variety of resources to obtain coordinate locations from the translated messages. Many of the volunteers were originally from Haiti and used their own geographical knowledge of the region combined with Open Street Map to pinpoint extract coordinates (Heinzelman and Waters, 2010). For the Ushahidi 2010 Gulf Oil Spill map we were unable to determine which specific methods were used to geo-locate implicit location information.

The software platform used by TtT extracts geospatial metadata using the Twitter API if such metadata exists. The software filters for location tags prescribed in the TtT micro-syntax or tags identifiable as spontaneously generated by the crowd that may include implicit location information, for example #loc or #lat and #long. These tags and the data after each tag were parsed into key-value pairs to populate the database. Location pairs, along with identified place names or event tags like Joplin or Boulder were geocoded using GeoKit (geokit.rubyforge.com), which can geocode textual information across a number of different geocoding services. Volunteers could review the resulting coordinate pairs, which were then entered into the database if approved.

COMPARING USHAHIDI AND TWEAK THE TWEET

Here we draw comparisons between Ushahidi and TtT in three key dimensions. First, we describe what types of data and variables are captured by each effort. Next, we compare the interactive mapping tools that each platform provides. We conclude our comparisons by characterizing how each platform has resulted in tangible actions by responders and emergency managers. We use four recent disasters in these comparisons. For Ushahidi, we explore its use in the 2010 Haiti earthquake and 2010 Gulf Oil Spill. For TtT we focus on the 2011 Joplin tornado and 2010 Boulder Fourmile Canyon fire. We were unable to find directly overlapping events for both platforms. Ushahidi deployments tend to focus on larger disasters rather than the localized events focused on by TtT. For one overlapping event, the 2010 Haiti earthquake, TtT had not yet implemented mapping tools, and in other overlapping events (like 2010 Pakistan floods) TtT has integrated their efforts with Ushahidi.

Field type	Ushahidi Field Name	TtT Field Name Boulder[1] Joplin[2]	Definition - TtT	Definition - Ushahidi
Record ID	#	Record ID[1] / ID[2]	Unique identifier	Unique identifier
Event		Event[1]	Event Hashtag – used for Place location	
Categorization	Category	Report Type[1,2]	Primary Hashtag – only one allowed – used for Key legend in Web Maps	Multiple categories allowed – used for Web Map Category Filter in Legend
Report	Incident Title	Report [2]	Partial parsed tweet with hashtags removed for pop-up display	Report Title for Web Report
Details		Details [1]	Partial parsed tweet with hashtags removed for pop-up display	
Original Report	Description	Text[2]	Original tweet	Original Message (in original language and translated if necessary)
Date/Time Stamp	Incident Date	Time[1,2]	Tweet time stamp	Message time stamp
Date_Time		Date_Time [2]	Time contained in tweet message	
Info		Info[1]	Volunteer added comment	
Source		Source[1,2]	Twitter user	
Contact		Contact[1,2]	Name, number, web page or other contact info contained in tweet	
Completed		Complete[1,2]	Indication if report was acted upon	
Status		Status[1]	? All N/A	
Verification	Verified	Verified[1]	? All N/A	Corroborated via incident report credibility vote
Actionable		Actionable[1]	? All N/A	
Approved	Approved			Map location approved
Author		Tweet Author[1] / Author[2]	The author of the record in the spreadsheet or author of retweet	
Tweet		Tweet[1]	Original Tweet	
Photo URL		Photo URL[1], Photo[2]	URL to photo	
Video		Video [2]	URL to Video	
Location (Text)	Location	Location[1,2]	Parsed location string	Parsed location string
Mapped		Mapped[1]	? All N/A	
Longitude	Longitude	GPS Long[1,2]	Derived Longitude	Derived Longitude
Latitude	Latitude	GPS Lat[1,2]	Derived Latitude	Derived Latitude

Table 2. Comparing data collected from TtT and Ushahidi

Raw Data

Data generated during a disaster from social media networks tend to be ephemeral and if it is not collected during the disaster, it can be difficult to conduct related research after the fact. Collecting raw data from Twitter older than two weeks has become challenging due to changes in the Twitter API that forbid certain types of archiving. Here, we conduct our analysis using the spreadsheets gathered from each application and additional analytical results from the PeopleBrowsr (www.peoplebrowsr.com) service which provides 1000 days of social media content and social analytics for marketers (not including SMS). We did not include PeopleBrowsr analytics for the 2010 Haiti earthquake because that data collected was primarily from SMS, and PeopleBrowsr analytics are not available for the 2010 Boulder Fourmile Canyon fire due to a small number of reports.

In Table 2 we list all fields we discovered in the TtT and Ushahidi spreadsheets and our interpretation of the definitions for each field type for each application. Common fields which we think share a common meaning across both platforms are highlighted. The Ushahidi platform has fewer fields (eight vs. twenty-five for TtT) and they do not vary between the two incidents. TtT has variation in field names and the number of fields. We note that it is difficult to differentiate between the terms Status, Actionable and Verified in the TtT fields.

In the content summary shown in Table 3 the Ushahidi field "Approved" always shows a rating of 100%. According to Ushahidi documentation all messages are "approved" once valid location coordinates are determined and an administrator approves the content. Reports that are not yet approved are not displayed. The "verified" field indicates a report is submitted by or corroborated by a trusted source or an administrator. Table 3 shows that 6% and 40% of the records were corroborated in the 2010 Haiti earthquake and 2010 Gulf Oil Spill events. Raw data from TtT did not reveal the meaning of the codes "Status", "Actionable", and "Verified."

The 2010 Gulf Oil Spill Ushahidi spreadsheet lists the first incident date eleven days after the Deepwater Horizon explosion. The total number of tweets with the #oilspill keyword from April 10th to October 18th, 2010 according to PeopleBrowsr, is 22,199. The Louisiana Bucket Brigade collected 2952 reports according to their spreadsheet, representing approximately 13% of the total Twitter traffic by the PeopleBrowsr estimate. Of note is that all 2952 reports were geo-located. Additionally, there were only 9 tweets on the day after the explosion and no tweets for the next 17 days. The traffic over six months highlights the extended nature of the disaster.

The 2011 Joplin tornado data starts the day after the tornado and ends 27 days after the tornado. According to PeopleBrowsr there were 333,387 total Twitter mentions of the #Joplin keyword from May 13 to June 13th, 2011. TtT identified 504 tweets that were entered into the spreadsheet. This represents approximately 0.02% of the total Twitter traffic if the PeopleBrowsr estimates are correct. This highlights the challenge associated with harvesting social media communications during temporally-limited crises. It is also interesting to note that 65% of the 504 records in the TtT spreadsheet for the 2011 Joplin tornado and 54% of the 522 records for the 2010 Boulder Fourmile Canyon fire included locations. Examination of the raw data reveals frequent status communication between volunteers that was not mapped because it was not relevant to the event itself.

	Incident	Incident Date	Reports	First Report Date	Last Report Date	%Verified	%Approved	%Actionable	%Complete	%LAT/LONG
TtT	Joplin Tornado[2]	5/22/2011 5:34 PM	504	5/23/2011 12:11 AM	6/13/2011 11:10 PM	N/A[1]		N/A[2]	0.4%[2]	65%
	Boulder Fourmile Fire[1]	9/6/2010 10:00 AM	522	9/8/2010 5:50 PM	9/17/2010 9:33 PM			N/A[1]	N/A[1]	54%
Ushahidi	Haiti Earthquake	1/12/2010 4:53 PM	3589	1/12/2010 4:08 AM	5/18/2010 4:26 PM	6%	100%			100%
	Gulf Oil Spill	4/10/2010 10:00 PM	2952	4/21/2010 1:44 PM	10/18/2010 10:07 PM	40%	100%			100%

Table 3. Summary of Ushahidi and TtT Spreadsheet Content

Maps

Cartographic representation of crisis mapping represents another challenge in the use of social media communications for disaster response because of the need to display large volumes of data while avoiding information overload. This is complicated further by the fact that potential users of crisis maps, including citizens, responders, volunteers, journalists and managers will have different expectations influenced by their social and physical relation to the crisis event (Liu and Palen, 2010). Field and O'Brien (2010) recognize that given the growth of social media communications and the geospatial component integral to an interconnected world, good cartography is crucial for creating maps with a purpose that are more than one-dimensional.

Crisis maps created by Ushahidi and TtT are quite similar (Figure 1) in terms of their core features. Both platforms utilize simple interactive map mashups and categorized point symbols to represent reports. Ushahidi has the ability to generalize dense sets of reports into aggregated symbols, making it scalable to larger datasets. Both platforms have recently introduced temporal displays to highlight report frequency over time (frequency graph in Ushahidi and time-categorized markers in TtT). The overall map and interface aesthetic is significantly more refined in current implementations of Ushahidi, perhaps reflecting its relative maturity compared to TtT.

Neither platform supports significant geospatial analysis capabilities. Basic filtering controls are available to winnow the dataset, but there are no quantitative spatial analysis methods available to identify clusters or to compare current patterns to past patterns. A significant difference between platforms is that Ushahidi provides alerting tools for users to "listen" for reports from a given area or matching a given set of thematic criteria.

Spatial data interoperability in both platforms is supported through spreadsheet downloads of raw data, making it possible for users to ingest collected information into a full-featured GIS if necessary.

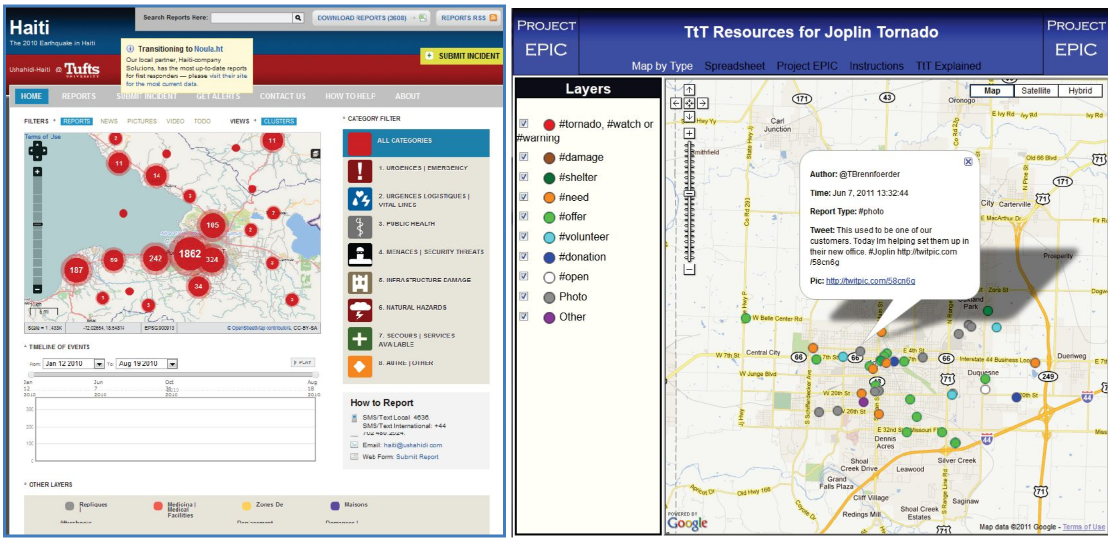

Figure 1. Examples of Ushahidi (left) and Tweak the Tweet (right) Mapping Interfaces

Evidence of Action

Understanding the effectiveness of efforts like TtT and Ushahidi is a difficult task. We concentrated our efforts on identifying impacts from media reports, after action reports, and TtT and Ushahidi's own assessments.

Project Epic and TtT gained recognition for their efforts during the 2010 Boulder Fourmile Canyon fire when CNN ran a story which credited TtT for integrating crisis information through the use of volunteers and social media (Spellman, 2010). One article cites Project Epic's use of a map with geo-located tweets to tracking the fire movements from citizen reports (Orlando, 2011). Another news report (Petty, 2010) describes how TtT was used to gather and map data from the 2010 Boulder Fourmile Canyon fire, including information not provided through official emergency response channels. We did not find many media reports of the specific use of TtT after the 2011 Joplin tornado, but the use of the syntax was promoted by organizations like Crisis Commons.

Following the 2010 Haiti earthquake, Craig Fugate, Director of FEMA, tweeted that the Ushahidi Haiti Map was the most "comprehensive and up-to-date map available to humanitarian organizations" (Heinzelman and Waters, 2010). Newsweek's profile of the Ushahidi efforts after Haiti indicate that the crisis map resulted in saving lives (Ramirez, 2010). One after action report (UN-SPIDER, 2010) notes that the US Marines used Ushahidi to coordinate locations and direct relief efforts. They also indicate that data from Ushahidi was used to direct Coast Guard responders for search and rescue. The Ushahidi Blog (blog.ushahidi.com) highlights multiple examples of action items and response efforts generated by Ushahidi, including food and water deliveries.

The 2010 Gulf Oil Spill Ushahidi map represented a different use of this technology in disaster response. The oil spill was not as much a direct threat to lives as it was a threat to local economies and the environment. Use of the map has been primarily to raise awareness of the ongoing ecological disaster and to document the damage. One article highlighted the fact that the information collected would be useful in long-term recovery efforts and could also be used in future legal actions over long-term damage to ecosystems and livelihoods (Sutter, 2010).

A common thread throughout the reports we reviewed is that the impact of geospatial social media platforms on tangible emergency response actions is not yet well-defined. While both have received media attention and have clearly captured public interest, there are few specific examples of the information leading to different decision-making patterns, widespread allocation of resources, or information leading to the rescue of disaster victims.

Future Design Considerations and Conclusions

The recent use of social media communications in disaster response is largely driven by volunteer organizations. Engagement with volunteers during the 2010 Haiti earthquake by TtT prompted TtT developers to focus research on a second layer of crowdsourcing: communication between volunteers and response organizations (Starbird, 2011). By working with crisis volunteer organizations TtT has continued to promote the use of the micro-syntax after a disaster. Spurred by lessons learned from deployments of crisis mapping efforts like Ushahidi, the Standby Task Force was recently formed to organize volunteers and provide a dedicated technical interface to the humanitarian community to assist in dealing with new sources of information like social media.

A key design consideration going forward is to ensure effective mechanisms for disseminating and sharing information between first responders and crisis managers. A UN report suggests that efforts like Ushahidi and TtT may contribute to information overload, citing that a large percentage of the information gathered was already in the hands of relief organizations on the ground (UN-SPIDER, 2010). This does not take into consideration the value of mapping the information and providing crisis managers a mechanism to identify clusters where relief organizations and first responders could concentrate their efforts.

We have outlined several reasons why extracting the location from social media is difficult, and further research is needed in this area, especially in the role of automatic geocoding and disambiguation of text descriptions of place. Ushahidi point out that from 40,000 Haiti-related SMS messages only 5% were mapped (Norheim-Hagtun and Meier, 2010), but that does not mean that the other 95% did not contain any geospatial information. To support analytical reasoning and geospatial analysis we should be able to uncover patterns that reference physical and cultural regions, types of landforms, directional information and topological relations in addition to basic point locations. A range of recent research focuses on the challenges of extracting actionable information from social media. Munro (2011) proposes models to systematically identify actionable items using trending categories or topics, subwords, and spatiotemporal clusters. MacEachren, et al. (2011) discuss how SensePlace2, a geovisual analytics application, includes a crawler application designed to systematically query the Twitter API based on crisis-relevant keywords and phrases. Vieweg, et al. (2010) have recently described an automated methodology to detect messages that will be useful for situational awareness.

Credibility and verification of information is another area that needs to be addressed in future research. One report indicates that search teams found a high proportion of SMS reports about trapped wounded victims turned out to be coming from families wanting to recover their dead relatives (Harvard Humanitarian Initiative, 2011). Some recent research has focused on identifying ways in which credibility might be automatically assessed in Tweets by evaluating message content, user profile details, and message propagation (Castillo, et al., 2011).

Finally, we must develop effective cartographic representation techniques to ensure the usability of web maps for crises. A particular challenge for crisis mapping is that there are a wide range of expectations and technical skills associated the diverse group of people that need to use crisis maps, including citizens, responders, volunteers, journalists and managers. Not all groups are equally equipped to evaluate the results of geographical analysis. Maps are likely to be seen as credible evidence, even when the underlying data is of unknown quality.

There is no doubt that the contribution of social media communications during disaster has shifted the paradigm of emergency response to include at least a one-way social media dialog from those most affected. Mapping social media content provides a way to gather and visualize information from what can arguably considered the true first responders - the affected citizens who are the first to assess the situation and request assistance through social media. Driven by volunteers and advances in web-based technology, the proliferation of this information has grown faster than the analytical capabilities of disaster management organizations and workflows. TtT has contributed a method to filter, automate and direct information from social media sources during a disaster and Ushahidi has proven to be an effective and widely adoptable platform for displaying geospatially-oriented social media communications. However, TtT and Ushahidi have only tackled simple location-related problems and provided only rudimentary situational awareness and mapping capabilities to visualize the social media communication stream. Future research must focus on applications that go beyond basic crowdsourcing to develop information collections, analytical tools, coordination of communications, and mapping visualization to support all phases of disaster management. Future platforms developed with the volunteer community in mind will need to incorporate social media as one piece of an overall strategy to support situational awareness and response and recovery featuring effective two-way communications with citizens through social media.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Homeland Security under Award Number: 2009-ST-061-CI0001. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies of the U.S. Department of Homeland Security.

REFERENCES

1. S. Vieweg, A.L. Hughes, K. Starbird, L. Palen, Microblogging during two natural hazards events: what twitter may contribute to situational awareness, in, ACM, Atlanta, Georgia, USA, 2010, pp. 1079-1088.
2. S.B. Liu, L. Palen, The new cartographers: crisis map mashups and the emergence of neogeographic practice, Cartography and Geographic Information Science, 37 (2010) 69-90.

3. A.M. MacEachren, A. Jaiswal, A.C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, J. Blanford, SensePlace2: Geotwitter Analytics Support for Situation Awareness, in: IEEE Conference on Visual Analytics Science and Technology, Providence, RI, 2011.
4. American Red Cross, Social Media in Disasters, in: http://www.redcross.org/www-files/Documents/pdf/SocialMediainDisasters.pdf, accessed on 10/15/2011
5. International Telecommunication Union, Key Global Telecom Indicators for the World Telecommunication Service Sector, in: www.itu.int/ITU-D/ict/statistics/at_glance/KeyTelecom.html, accessed on 10/15/2011
6. K.M. Anderson, A. Schram, Design and implementation of a data analytics infrastructure in support of crisis informatics research (NIER track), in, ACM, Waikiki, Honolulu, HI, USA, 2011, pp. 844-847.
7. Harvard Humanitarian Initiative, Disaster Relief 2.0: The future of information sharing in humanitarian emergencies, in, 2011.
8. K. Starbird, Digital Volunteerism During Disaster: Crowdsourcing Information Processing, in: CHI '11 Workshop on Crowdsourcing and Human Computation, Vancouver, BC, 2011.
9. K. Starbird, L. Palen, "Voluntweeters": self-organizing by digital volunteers in times of crisis, in, ACM, Vancouver, BC, Canada, 2011, pp. 1071-1080.
10. O. Okolloh, Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information, Participatory Learning and Action, 59 (2009) 65-70.
11. P. Currion, C. de Silva, B. Van De Walle, Open source software for disaster management, Communications of the ACM, 50 (2007) 61-65.
12. M. Zappavigna, Ambient affiliation: A linguistic perspective on Twitter, New Media & Society, 13 (2011) 788-806.
13. L. Palen, K.M. Anderson, G. Mark, J. Martin, D. Sicker, M. Palmer, D. Grunwald, A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters, in: ACM-BCS Visions of Computer Science Conference, Edinburgh, UK, 2010, pp. 1-12.
14. J. Heinzelman, C. Waters, Crowdsourcing Crisis Information in Disaster-Affected Haiti, United States Institute of Peace, 252 (2010) 1-16.
15. S. Dosemagen, Ushahidi Used to Create Oil Spill Crisis Map, in: The Ushahidi Blog, 2010.
16. Ushahidi, Ushahidi API, in: http://wiki.ushahidi.com/doku.php?id=ushahidi_api, accessed on 10/15/2011
17. P. Meier, R. Munro, The Unprecedented Role of SMS in Disaster Response: Learning from Haiti, SAIS Review, 30 (2010) 91-103.
18. R. Munro, Subword and spatiotemporal models for identifying actionable information in Haitian Kreyol, in, Association for Computational Linguistics, Portland, Oregon, 2011, pp. 68-77.
19. A. Bellucci, A. Malizia, P. Diaz, I. Aedo, Framing the design space for novel crisis-related mashups: the eStoryS example, in: Information Systems for Crisis Management and Response (ISCRAM 2010), Seattle, WA, 2010, pp. 1-10.
20. K. Field, J. O'Brien, Cartoblography: experiments in using and organising the spatial context of micro-blogging, Transactions in GIS, 14 (2010) 5-23.
21. J. Spellman, Heading off disaster, one tweet at a time, in: CNN Tech, www.cnn.com/2010/TECH/social.media/09/22/natural.disasters.social.media/index.html,accessed on 10/15/2011
22. J. Orlando, Turning Disaster Response On Its Head, in: Continuity Insights, http://www.continuityinsights.com/articles/turning-disaster-response-on-its-head, accessed on 10/15/2011
23. D. Petty, Evacuees use social media to keep up on Boulder wildfire disaster developments, in: Denver Post, MediaNews, Denver, CO, 2010.
24. J. Ramirez, 'Ushahidi' Technology Saves Lives in Haiti and Chile, in: Newsweek, 2010.
25. UN-SPIDER, Lessons from Haiti, Coordinates, 6 (2010) 27-31.
26. J.D. Sutter, Citizens monitor Gulf Coast after oil spill, in: CNN Tech, articles.cnn.com/2010-05-06/tech/crowdsource.gulf.oil_1_oil-spill-gulf-coast-jeffrey-warren, accessed on 10/15/2011
27. I. Norheim-Hagtun, P. Meier, Crowdsourcing for Crisis Mapping in Haiti, innovations, 5 (2010) 81-89.
28. C. Castillo, M. Mendoza, B. Poblete, Information credibility on twitter, in, ACM, Hyderabad, India, 2011, pp. 675-684.