An Emotional Step Towards Automated Trust Detection in Crisis Social Media

Shane E. Halse
Pennsylvania State University
seh297@ist.psu.edu

Andrea Tapia
Pennsylvania State University
atapia@ist.psu.edu

Anna Squicciarini
Pennsylvania State University
asquicciarini@ist.psu.edu

Cornelia Caragea
University of North Texas
Cornelia.Caragea@unt.edu

ABSTRACT

To this date, research on crisis informatics has focused on the detection of trust in Twitter data through the use of message structure, sentiment, propagation and author. Little research has examined the effects of perceived emotion of these messages in the crisis response domain. Toward detecting useful messages in case of crisis, we examine perceived emotions of these messages and how the different emotions affect the perceived usefulness and trustworthiness. Our analysis is carried out on two datasets gathered from Twitter concerning hurricane Sandy in 2012 and the Boston Bombing 2013. The results indicate that there is a significant difference in the perceived emotions that contribute towards the perceived trustworthiness and usefulness. This could have impacts on how messages from social media data are analyzed for use in crisis response.

Keywords

Twitter. Sandy. Hurricane. Boston. Bombing. Trust. Usefulness. Sentiment. Emotion.

INTRODUCTION

We believe that data directly contributed by citizens, and data scraped from bystanders witnessing a disaster, have strongly positive potential to give responders more accurate and timely information than is possible with traditional information gathering methods. Many emergency decision makers see the data produced through crowd sourcing and social media as ubiquitous, rapid and accessible—with the potential to contribute to situational awareness (Vieweg et al. 2010). In response to increased online public engagement and the emergence of digital volunteers, professional emergency responders have sought to better understand how they too can use online media to communicate with the public and collect intelligence (Denef, Augustin, Bayerl, & Kaptein, 2013; Latonero & Shklovski, 2011; Sutton et al., 2015).

Due to the perceived lack of authentication and validation of content posted in Twitter micro-blogs, large-scale responders have been reluctant to incorporate social media data into the process of assessing a disaster situation and the subsequent decision-making process to send aid workers and supplies to disaster locations. Committing to the mobilization of valuable and time sensitive relief supplies and personnel, based on what may turn out be illegitimate claims, has been perceived to be too great a risk. Incorporating the products of digital volunteer activity into professional emergency practice has proved to be challenging due to issues with credibility, liability, training, and organizational process and procedure (Hughes and Palen 2012; Starbird and Palen 2013).

Our study resulted in a wealth of useful and informative insights. In particular, our analysis allows us to take a first step toward understanding whether emotions used during a tweet will affect the outcome of community vetted information in terms of tweet trustworthiness and usefulness. We focused on addressing the following research questions: For two different disaster types, man-made and natural, what are the effects of perceived emotions on trustworthiness and usefulness?

Our findings indicate that there is an effect of emotion type, particularly fear, sadness, and neutral emotions on the trustworthiness and usefulness of a Tweet.

RELEVANT LITERATURE

Using social media feeds as information sources during a large-scale event is highly problematic for several reasons, including the inability to verify either the person or the information that the person posts (Tapia, Bajpai, Jansen, Yen, & Giles, 2011; Tapia, Moore, & Johnson, 2013; Starbird & Paylen, 2013).

The research involving veracity in technologically mediated environments has had two distinct approaches. The first approach looks at the person supplying the information, while the second looks at the information itself. Identifying who is providing information, whether the person is reliable, credible and in a position to know information are extremely valuable factors to establish trustworthiness (Grabner-Kräuter, Kaluscha, & Fladnitzer, 2006). On a personal level, analysis of sentiment, or the implied emotional state of a microblogger has been proven useful in political debate analyses, earthquakes, and during national security incidents (Diakopoulos & Shamma, 2010; Qu, Huang, Zhang, & Zhang, 2011).

In recent work by (Gupta, Kumaraguru, & Castillo, 2014), a real-time system, called *TweetCred*, was developed to assign a credibility score to tweets in a user's timeline; however, it does not take the detection of emotions into account. Their findings indicate that the system assigned a trust score lower than expected by their participants which could be explained by the perceived emotional context. Using features extracted from users' posting behavior and tweets' social context, the newsworthy events were classified as credible or not credible (Castillo, Mendoza, & Poblete, 2011). Their work tried to model whether end-users would believe the information reported in tweets is true or not, but was not concerned with detecting whether the information in tweets was itself accurate or contained an emotion. (Dailey & Starbird, 2014) explored techniques such as *visible skepticism* to help control the spread of false rumors, but did not intend to automatically detect false rumors. Most research in disaster-related area has been performed post-hoc, and the most important aspect of any intelligence received, intelligence that is actionable and precisely geo-located, has not yet been achieved and is also complicated by translation and language understanding (McClendon & Robinson, 2012; Munro, 2011).

METHODS

In 2014 we started the complex process of creating gold-standard datasets of disaster-related data. Given our collections of tweets related to the Hurricane Sandy and Boston bombings, we derived two sets of rules. These rules included labeling events as relevant, i.e., disaster or non-disaster events, trustworthy and their usefulness to first-responders. Based on the derived rules, we used Mechanical Turk (MTurk) to manually label the tweets. Utilizing crowd sourced data such Amazon Turk is representative of the volunteers that could be found during a crisis and is supported by the works by (Dailey & Starbird, 2014) and (Ann, Denis, & Hughes, 2012). In total about 3711 unique tweets were labeled (some had multiple labelers).

Two main dependent variables were investigated, the perceived trustworthiness and usefulness of each tweet. Each variable has 5 levels as depicted in table 1 and table 2. The participants were shown a tweet along with 9 questions asking about their perceptions of that tweet as seen in table 3. For the perceived trustworthiness users were asked "How trustworthy would you rate this tweet?" Likewise for the perceived usefulness users where asked "How useful would this tweet be to first responders (i.e. units trying to provide help)?"

Table 1 - Trustworthiness Coding Schema

Trustworthiness Coding	
Category	Description
DefT	Definitely Trustworthy
MaybeT	Maybe Trustworthy
MaybeUnt	Maybe Untrustworthy
MostCUnt	Most Certainly Untrustworthy
UnkT	Unknown

Table 2 - Usefulness Coding Schema

Table 2 - Usefulness Coding Schema

Usefulness Coding	
Category	Description
VeryUse	Very Useful
MaybeUse	Maybe Useful
MaybeNotUse	Maybe not Useful
NotUse	Not Useful
IDKUse	I don't know/Not applicable

In total we focused on the independent variable related to the emotion expressed within the tweet. This question was designed as an 8 point categorical and the response text and codes can be seen in table 3.

Table 3 - Independent Variable Coding Schema

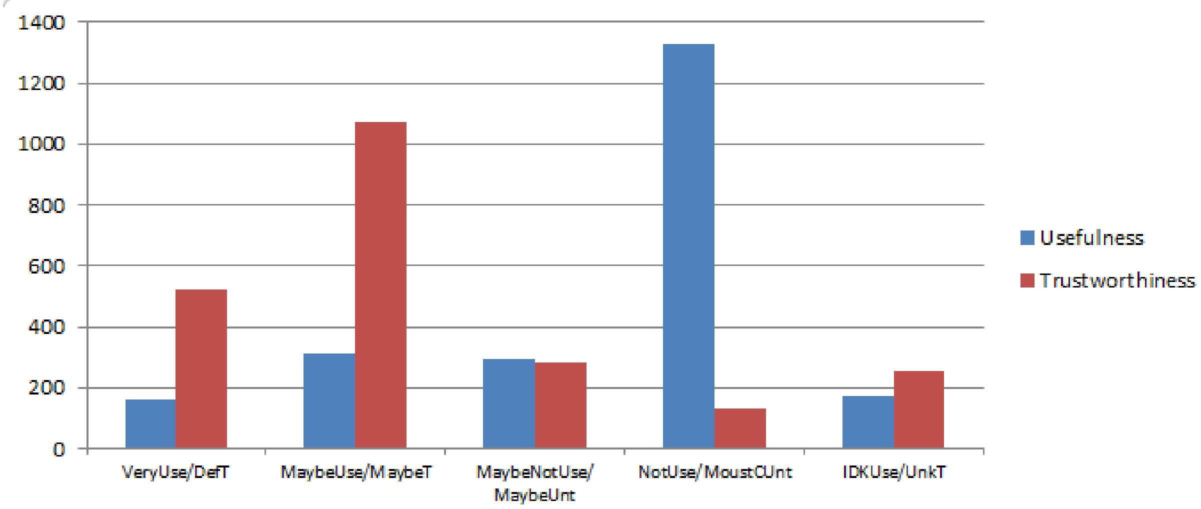
Variable	Description	Response Codes
EmoType	Which emotions does this tweet express?	anger, disgust, fear, happiness, sadness, surprise, neutral, irrelevant

Data description

The first data set used in our experiment was collected from Twitter during the disastrous Hurricane Sandy. Specifically, the dataset contains 12,933,053 tweets crawled between 10-26-2012 and 11-12-2012 using the hashtag sandy and hurricane. We randomly sampled a subset of 1711 tweets from the crawled data for our labeling tasks. The second data set used in our experiment was collected from Twitter using the hashtag prayforboston, Boston, bomb, and bombs during the Boston Bombing incident between 04-15-2013 and 04-25-2013 and contained 23,642,905 tweets. We randomly sampled a subset of 898 tweets from the crawled data for our labeling tasks.

Overall for the Hurricane Sandy dataset, we employed 1702 workers, who labeled 1711 unique tweets, each worker labeled multiple tweets and most tweets had multiple labelers. Of these, we disregarded 149 (8.71 %) tweet messages and their corresponding labelers data as these were perceived as not being related to Hurricane Sandy. For the Boston bombing dataset a total of 898 unique tweets were labeled (some had multiple labelers). Of these, we disregarded 207 (23.05%) tweet messages and their corresponding labels data as these were perceived as not being related to the Boston Bombing. Note as our dataset for the Boston bombing contained the filter "Boston" this resulted in a large number of tweets that were unrelated to the bombing. This allowed us to have tweets that were considered relevant to the Sandy Hurricane and the Boston Bombing by the labelers.

The following graph represents the distribution of the labeler's responses to each tweet with respect to the tweet being trustworthy and/or useful:



Statistical techniques

Multiple tests where completed in order to fully understand and describe the data. In order to prevent Type I error a large sample size was used for each test in addition to setting the significance level at $\rho < 0.05$. In order to answer the question of "For two different disaster types, man-made and natural, what are the effects of perceived emotions on trustworthiness and usefulness?" the following hypotheses was developed:

> H_1: There is a significant effect of perceived emotions on the perceived trustworthiness and usefulness rating found in both man-made and natural disaster data

This hypothesis would provide sustenance to the idea of adding the factor of perceived emotion type to tools developed to automate the detection of trust and usefulness.

A multinomial logistic regression was used to identify the relationships between the dependent variables of trustworthiness and usefulness with the independent variable of Emotion type. As we were primarily interested in the analysis within the disaster response domain, data with a response of "No" to the "is this tweet about the disaster", were removed. The logistic regression model is shown as the following form:

$$\ln\left(\frac{\rho}{1-\rho}\right) = \beta_0 + \beta_i X_i$$

where ρ is the probability of effects on the perceived trustworthiness and usefulness, $(\frac{\rho}{1-\rho})$is the odds of the effects on the perceived trustworthiness and usefulness; β_0 is a constant X_i is the vector of independent variables and β_i is the parameter for the ith independent variable. This was done twice, once for each of the datasets used within the analysis for this paper. A similar analysis can be found in (Hyun & Ditton, 2006)

Multinomial Logistic Regression

A multinomial regression analysis was used with the value of 0 (most certainly untrustworthy and not useful) selected as the reference point. The results are shown below.

Sandy data

The first analysis for the multinomial logistic regression was for the hurricane Sandy dataset's usefulness score the shown in Table 4 - Multinomial Logistic Regression (Sandy - Emotype vs Usefulness) and was found to be statistically significant ($\chi^2 = 345.97$; $\rho < 0.001$). It shows that two highest odds ratios are for the emotion of sadness and fear. That is, a tweet is perceived to have a sadness quality then it is 5.896 times more likely to be considered very useful than not at all useful.

Table 4 - Multinomial Logistic Regression (Sandy - Emotype vs Usefulness)

Multinomial Logistic Regression showing significant effects of Emotion type on Perceived Usefulness (Final model only included significant variables at .05 level)

How useful is the tweet?		Estimate	SE	t-value	p(Sig.)	Odds Ratio
Very Useful	(intercept)	-3.455	0.339	-10.206	0.000	0.032
	EmoTypefear	1.373	0.378	3.633	0.000	3.946
	EmoTypeneutral	1.006	0.355	2.834	0.005	2.734
	EmoTypesadness	1.775	0.421	4.218	0.000	5.896
Maybe Useful	(intercept)	-2.097	0.179	-11.709	0.000	0.123
	EmoTypefear	1.238	0.206	6.004	0.000	3.450
	EmoTypeneutral	0.718	0.191	3.756	0.000	2.050
	EmoTypesadness	1.404	0.248	5.662	0.000	4.071
	EmoTypesurprise	0.608	0.242	2.508	0.012	1.836
Maybe not Useful	(intercept)	-2.434	0.209	-11.667	0.000	0.088
	EmoTypefear	1.151	0.241	4.784	0.000	3.161
	EmoTypehappiness	0.592	0.249	2.374	0.018	1.808
	EmoTypeneutral	0.602	0.224	2.689	0.007	1.825
	EmoTypesurprise	1.176	0.256	4.589	0.000	3.241
IDK Useful	(intercept)	-3.013	0.274	-11.008	0.000	0.049
	EmoTypeirrelevant	1.317	0.296	4.456	0.000	3.733

Model chi-squared = 345.97; ρ<0.0001, log likelihood = -4896.3, Pseudo R^2 (McFadden) = 0.034124. The reference category is NotUse (Not Useful)

The second analysis was for the hurricane Sandy dataset's trustworthiness score the shown in Table 5 and was found to be statistically significant ($\chi^2 = 497.28$; $\rho < 0.001$). Note that two highest odds ratios are for the emotion of sadness and fear. That is, for example, a tweet is perceived to have a fear emotion then it is 10.354 times more likely to be considered very trustworthy than not trustworthy.

Table 5 - Multinomial Logistic Regression (Sandy - Emotype vs Trustworthiness

Multinomial Logistic Regression showing significant effects of Emotion type on Perceived trustworthiness (Final model only included significant variables at .05 level)

How trustworthy is the tweet?		Est.	SE	t-value	p(Sig)	Odds Ratio
Definitely Trustworthy	EmoTypefear	2.337	0.322	7.266	0.000	10.354
	EmoTypehappiness	0.541	0.250	2.167	0.030	1.718
	EmoTypeirrelevant	-0.790	0.243	-3.245	0.001	0.454
	EmoTypeneutral	1.231	0.215	5.723	0.000	3.426
	EmoTypesadness	1.686	0.390	4.320	0.000	5.400
Maybe Trustworthy	(intercept)	0.807	0.164	4.930	0.000	2.241
	EmoTypefear	2.037	0.305	6.684	0.000	7.670
	EmoTypeirrelevant	-0.557	0.202	-2.759	0.006	0.573
	EmoTypeneutral	0.959	0.192	5.000	0.000	2.609
	EmoTypesadness	1.309	0.373	3.514	0.000	3.704
	EmoTypesurprise	0.790	0.256	3.084	0.002	2.203
Maybe Untrustworthy	EmoTypefear	1.377	0.332	4.145	0.000	3.964
	EmoTypehappiness	0.530	0.247	2.151	0.032	1.438
Unkown Trustworthiness	EmoTypefear	0.886	0.339	2.615	0.009	2.424
	EmoTypeneutral	0.658	0.211	3.118	0.002	1.931

Model chi-squared = 497.28; ρ<0.0001, log likelihood = -6877.6, Pseudo R^2 (McFadden) = 0.034891. The reference category is MostCUnt (Most Certainly Untrustworthy)

Boston data

The third analysis for the multinomial logistic regression was for the Boston bombing dataset's usefulness score the shown in table 6 and was found to be statistically significant ($\chi2 = 150.04$; $\rho < 0.001$). It shows that highest odds ratio is that of the emotion fear meaning it is 9.419 times more likely to be considered very useful than not at all useful.

Table 6 - Multinomial Logistic Regression (Boston - Emotype vs Usefulness)

Multinomial Logistic Regression showing significant effects of Emotion type on Perceived Usefulness (Final model only included significant variables at .05 level)

How useful is the tweet?		Estimate	SE	t-value	p(Sig.)	Odds Ratio
Very Useful	(intercept)	-2.841	0.364	-7.810	0.000	0.058
	EmoTypefear	2.243	0.450	4.981	0.000	9.419
	EmoTypeneutral	0.862	0.384	2.247	0.025	2.367
	EmoTypesurprise	0.995	0.478	2.080	0.038	2.704
Maybe Useful	(intercept)	-2.522	0.313	-8.048	0.000	0.080
	EmoTypefear	2.165	0.399	5.432	0.000	8.718
	EmoTypeirrelevant	0.991	0.351	2.818	0.005	2.693
	EmoTypeneutral	1.387	0.325	4.268	0.000	4.002
Maybe not Useful	(intercept)	-2.147	0.264	-8.128	0.000	0.117
	EmoTypefear	0.943	0.422	2.235	0.025	2.569
	EmoTypeneutral	0.912	0.279	3.272	0.001	2.490
	EmoTypesurprise	0.812	0.365	2.228	0.026	2.253
IDK Useful	(intercept)	-2.617	0.328	-7.990	0.000	0.073
	EmoTypehappiness	0.889	0.411	2.161	0.031	2.433
	EmoTypeirrelevant	0.798	0.374	2.137	0.033	2.222
	EmoTypeneutral	0.772	0.347	2.225	0.026	2.164

Model chi-squared = 150.4; ρ<0.0001, log likelihood = -2720, Pseudo R^2 (McFadden) = 0.026903. The reference category is NotUse (Not Useful)

The fourth analysis was for the Boston bombing dataset's trustworthiness score the shown in and was found to be statistically significant ($\chi2 = 194.71$; $\rho < 0.001$). It shows that the highest odds ratios are for the emotion of sadness. This means that an emotion of sadness during a man-made disaster would lead to a higher trustworthiness score.

Table 7 - Multinomial Logistic Regression (Boston - Emotype vs Trustworthiness)

Multinomial Logistic Regression showing significant effects of Emotion type on Perceived trustworthiness (Final model only included significant variables at .05 level)

How trustworthy is the tweet?		Estimate	SE	t-value	p(Sig.)	Odds Ratio
Definitely Trustworthy	EmoTypefear	2.010	0.678	2.967	0.003	7.467
	EmoTypehappiness	1.550	0.469	3.305	0.001	4.711
	EmoTypeneutral	2.069	0.358	5.784	0.000	7.917
	EmoTypesadness	2.241	0.601	3.728	0.000	9.400
	EmoTypesurprise	1.131	0.497	2.275	0.023	3.100
Maybe Trustworthy	(intercept)	1.322	0.252	5.252	0.000	3.750
	EmoTypefear	1.408	0.647	2.177	0.029	4.089
	EmoTypeneutral	1.612	0.316	5.106	0.000	5.011
	EmoTypesadness	1.711	0.570	2.999	0.003	5.533
Maybe Untrustworthy	(intercept)	0.615	0.278	2.217	0.027	1.850
	EmoTypeirrelevant	-0.766	0.347	- 2.209	0.027	0.465
	EmoTypeneutral	0.762	0.347	2.199	0.028	2.144
Unkown Trustworthiness	EmoTypeneutral	0.864	0.369	2.342	0.019	2.372
	EmoTypefear	0.886	0.339	2.615	0.009	2.400

Model chi-squared = 194.71; ρ<0.0001, log likelihood = -2985.4, Pseudo R^2 (McFadden) = 0.03158. The reference category is MostCUnt (Most Certainly Untrustworthy)

In order to investigate the significant effects of emotion types on our trustworthiness and usefulness score we have prepared the following summary table below:

Table 8 - Summary table indicating the significant emotion types on usefulness and trustworthiness with α < 0.05

Emotion Type	Sandy Data (Trustworthiness)				Sandy Data (Usefulness)				Boston Data (Trustworthiness)				BostonData (Usefulness)			
	VT	MT	MNT	IDKT	VU	MU	MNU	IDKU	VT	MT	MNT	IDKT	VU	MU	MNU	IDKU
anger																
disgust																
fear	10.354	7.67	3.964	2.424	3.946	3.45	3.161		7.467	4.089		2.4	9.419	8.718	2.569	
happiness	1.718		1.438				1.808		4.711							2.433
sadness	5.4	3.704			5.896	4.071			9.4	5.533						
surprise		2.203				1.836	3.241		3.1				2.704		2.253	
neutral	3.426	2.609		1.1931	2.734	2.05	1.825		7.917	5.011	2.144	2.372	2.367	4.002	2.49	2.164
irrelevant	0.454	0.573						3.733			0.465			2.693		2.222

VT = Very rustworthy, MT = aybe Trustworthy, MNT = Maybe not trustworthy, IDKT = I don't know trustworthy
VU = Very useful, MU = maybe useful, MNU = maybe not useful, IDKU = I don't know useful

As we can see from the table (Table 8 - Summary table indicating the significant emotion types on usefulness and trustworthiness with α < 0.05 when using a multinomial logistic regression to compare the significant likelihoods that a user will select an option other than Most certainly untrustworthy we can see the following patterns;

- If a tweet has a perceived emotion type of anger or disgust, then there is no significant difference in the selection of the perceived trustworthiness or usefulness.
- If a tweet has a perceived emotion type of happiness, sadness, surprise or irrelevant then there is a partial significant difference in the selection of the perceived trustworthiness or usefulness.
- The most interesting finding is that of the emotion type fear and neutral. Both of these emotion types can significantly affect the likelihood of selecting an option other than most certainly untrustworthy.

DISCUSSION

It is important to remember that the barriers of social media data adoption are broad and numerous. Despite this, the advantages and potential information that can be found within the network fair outweighs the difficulty of the work that must be performs in using it. (Tapia et al., 2011) have already described pockets of use of social media data and illustrate both the frustration and hope of one-day being able to use this data effectively. As such, automatically synthesizing of this data into a manageable quantity should be the goal of anyone hoping to overcome the barriers to it adoption. While previous research has begun to attempt to do this through the avenue of investigating trust and how to automate its assessment, this research shows that the addition of detecting a tweet's perceived emotion can further validate the trust and usefulness measure.

Our analysis focuses primarily on the type of emotion used during a tweet as this was deemed to be the most interesting and novel factor and serves as the main driver for the multinomial logistic regression. As seen in the summary table (Table 8) the two primary emotions that affected the selection of the trust score value were fear and neutral (no emotion). The first emotion type, fear, could be explained by people sympathizing with the tweet authors. That is when someone is scared or sad, human nature is to sympathize and try to help; this in turn would lead one to be more likely to believe (trust) the data being presented. In addition the neutral emotion could be attributed to Tweet presenting facts or information in which the data is very cut and dry. This could be an explanation as to why the neutral emotion type played a smaller role (than that of fear) in the selection of the trust score. Stepping back, one can see that the H_1 should be accepted as the results support emotion types affecting perceived usefulness and trustworthiness rating.

CONCLUSIONS

To fully develop an automated trust score for Twitter tweets is no easy however this research has shown an additional technique for generating this score. That is, by investigating methods to detect the emotions found within a Twitter message, in particular, fear and neutral emotion would prove beneficial in providing the crisis responders an accurate assessment of the trust of the message. Further as social media data provide a plethora of data, tools that would generate a trust assessment based on both man-made and natural disasters should use detection of emotion type (possibly through sentiment analysis) to improve the automation of the trust assessment of social media data. That is, by using the detection of emotion a tweet that is deemed highly trustworthy or useful with a high fear level, we could flag it for further vetting. Tweets with a high trustworthy and usefulness rating and a neutral emotion could be passed through as more informational and further examined.

In utilizing these tools the barrier of the amount of data found can be overcome. That is, rather than have valuable resources devoted to sifting through and analyzing every tweet during a crisis; tools could partially automate this tedious process. In doing this these tools would be able provide more relevant information to first responders. In detecting emotions in the tweet (particularly fear) the tools could mark the message for further vetting. This would prove useful in situations where a Tweet has been deemed untrustworthy through other methods, and yet, people still want to believe it. In turn, this would allow the crisis response managers to have a better understanding of the data leading to better situational awareness, and thus allowing them to make effective decisions in time critical situations.

REFERENCES

1. Ann, L., Denis, S., & Hughes, A. L. (2012). Trial by Fire : The Deployment of Trusted Digital Volunteers in the 2011 Shadow Lake Fire, (April), 1–10.
2. Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on twitter. *Proceedings of the 20th International Conference on World Wide Web - WWW '11*, 675. http://doi.org/10.1145/1963405.1963500
3. Cheong, M., & Lee, V. C. S. (2011). A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via Twitter. *Information Systems Frontiers*, *13*(1), 45–59. http://doi.org/10.1007/s10796-010-9273-x
4. Dailey, D., & Starbird, K. (2014). Visibl e Skepticism : Community Vetting after Hurricane Irene, (May), 775–779.
5. Denef, S., Augustin, S., Bayerl, P. S., & Kaptein, N. (2013). Social Media and the Police—Tweeting Practices of British Police Forces during the August 2011 Riots, (August 2011), 3471–3480.
6. Diakopoulos, N. a, & Shamma, D. a. (2010). Characterizing debate performance via aggregated twitter sentiment. *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1195. http://doi.org/10.1145/1753326.1753504
7. Grabner-Kräuter, S., Kaluscha, E. A., & Fladnitzer, M. (2006). Perspectives of online trust and similar constructs. In *Proceedings of the 8th international conference on Electronic commerce The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet - ICEC '06* (p. 235). New York, New York, USA: ACM Press. http://doi.org/10.1145/1151454.1151496
8. Gupta, A., Kumaraguru, P., & Castillo, C. (2014). TweetCred : Real-Time Credibility Assessment, 228–243.
9. Hyun, W., & Ditton, R. (2006). Using multinomial logistic regression analysis to understand anglers willingness to substitute other fishing locations. *Burns R, Robinson K (eds) Proceedings of the …*, 248–255. Retrieved from http://www.nrs.fs.fed.us/pubs/gtr/gtr_nrs-P-14.pdf#page=257
10. Latonero, M., & Shklovski, I. (2011). Emergency Management , Twitter , and Social Media Evangelism, *3*(December), 1–16. http://doi.org/10.4018/jiscrm.2011100101
11. McClendon, S., & Robinson, A. C. (2012). Leveraging Geospatially-Oriented Social Media Communications in Disaster Response. *Iscram*, (April), 2–11. http://doi.org/10.4018/jiscrm.2013010102
12. Munro, R. (2011). Subword and spatiotemporal models for identifying actionable information in Haitian Kreyol, (June), 68–77.
13. Qu, Y., Huang, C., Zhang, P., & Zhang, J. (2011). Microblogging after a major disaster in China: a case study of the 2010 Yushu earthquake. *… of the ACM 2011 Conference on …*, 25–34. http://doi.org/10.1145/1958824.1958830

14. Starbird, K., & Paylen, L. (2013). Working & Sustaining the Virtual " Disaster Desk ."
15. Sutton, J., Gibson, C. Ben, Spiro, E. S., League, C., Fitzhugh, S. M., & Butts, C. T. (2015). What it Takes to Get Passed On: Message Content, Style, and Structure as Predictors of Retransmission in the Boston Marathon Bombing Response. *Plos One*, *10*(8), e0134452. http://doi.org/10.1371/journal.pone.0134452
16. Tapia, A. H., Bajpai, K., Jansen, J., Yen, J., & Giles, L. (2011). Seeking the Trustworthy Tweet: Can Microblogged Data Fit the Information Needs of Disaster Response and Humanitarian Relief Organizations. *Proceedings of the 8th International ISCRAM Conference*, (May), 1–10. Retrieved from http://irevolution.files.wordpress.com/2011/06/tapia-seeking_the_trustworthy_tweet-161_a.pdf
17. Tapia, A. H., Moore, K. a, & Johnson, N. (2013). Beyond the Trustworthy Tweet: A Deeper Understanding of Microblogged Data Use by Disaster Response and Humanitarian Relief Organizations. *Proceedings of the 10th International ISCRAM Conference*, (May), 770–779. Retrieved from http://star-tides.net/sites/default/files/documents/files/Beyond the Trustworthy Tweet - A Deeper Understanding of Microblogged Data Use by Disaster Response and Humanitarian Relief Organizations.pdf