

Steve Schifferes, Nic Newman, Neil Thurman, David Corney, Ayse Goker, and Carlos Martin

IDENTIFYING AND VERIFYING NEWS THROUGH SOCIAL MEDIA: Developing a user-centred tool for professional journalists

Identifying and verifying new information quickly is the key issue for journalists who use social media. This paper examines what tools journalists think they need to cope with the growing volume and complexity of news on social media, and what improvements are needed in existing systems. It gives some initial results from a major EU research project (Social Sensor), involving computer scientists, journalists, and media researchers, that is designing a new tool to search across social media for news stories, to surface trends, and to help with verification. Preliminary results suggest that an effective tool should focus on the role of key influencers, and should be customisable to suit the particular needs of individual journalists and news organisations.

Keywords: computer aided reporting, journalists, social media, sourcing, Twitter, verification

Introduction

As social media has become an important source of news for journalists, there are increasing concerns about the nature and reliability of news on networks like Twitter and Facebook. Journalists are looking to quickly identify relevant information and trusted sources, but how can they also spot misinformation, faked pictures, and the setting of false trails? The monitoring of social media has become a time-consuming extra task whilst the problematic nature of bespoke verification and the requirement to clear rights has reduced the ability of journalists to make full use of social media. It follows that without tackling these thorny issues the use of social media as a source of news will be increasingly problematic. This article is based on research carried out for the EU Social Sensor project, a research consortium¹ of 10 institutions across Europe (including Yahoo, IBM, and Deutsche Welle) which has been tasked with developing new software tools that can help journalists in their work with social media. It will be argued that a multi-faceted approach to the surfacing of important trends and to measuring trust and reliability is essential in developing any new tool. Furthermore, any verification tool must be capable of adjustment to the perceived needs of journalists (which may vary in different news organisations). The approach of the Social Sensor research project is user-centred, based on the perceived needs of professional journalists working in a real-time environment.

Literature Review

There is a growing body of literature—in both the fields of media and communications, and information science research—on how journalists and the public use social media to find news. Recent research (Reuters Institute 2013; Pew

2012) has shown rapidly increasing use of social media to discover news stories. The importance of social media both as a source of news and means of dissemination is growing for news organisations (Newman 2009, 2011; Newman, Dutton, and Blank 2012; Hahn 2013; Lasorsa, Lewis, and Holton 2012) with Twitter and Facebook becoming particularly important.

Researchers in computer science have also begun to look at the relationship between social media (particularly Twitter) and the news. One of the earliest papers (Kwak et al. 2010) demonstrated that news on Twitter closely tracks the mainstream news agenda, highlighted the importance of retweets, and suggested the timeframe for discussion of any topic was short. Castillo et al. (2012) showed that the importance of a news story on Twitter became apparent very quickly; their research showed that it was possible to predict the overall traffic that a news article will receive with 10 minutes of measuring reactions to it on social media. The relationship between the news agenda and Twitter has also been studied by Petrovic et al. (2013) who reported that Twitter tracked mainstream newswires in finding breaking news stories.

The role of influential experts in spotting and verifying the news has been the subject of considerable investigation. In looking at the Chile earthquake in 2010, Castillo, Mendoza, and Poblete (in press) found that the identification of influential tweeters, and certain characteristics of tweets (such as their length, and whether they contained links or pictures), could help distinguish between true and false reports concerning the natural disaster, although the authors commented that computer algorithms alone could not give a 100% guarantee of accuracy.

In further work, Diakopoulos, De Choudhury, and Naaman (2012) tried to create a customisable tool where journalists could collectively assess the validity of social media sources (using the Tottenham Riots in 2011 as the data source), using visual icons to expose relevant information derived from Twitter, such as the likely location of a posting. They concluded that journalists needed to use a variety of methods for rating the usefulness of tweets. The importance of crowd-sourced news recommendations was emphasised by Morales, Gionis, and Lucchese (2012) who found that a more sophisticated analysis of recommendations and popularity on social networks could help predict major stories.

Further refinements emerged from the work of Hsieh and Lehmann. Hsieh et al. (2013) showed that crowd-sourcing could identify key topics, but that experts could enhance the process. They also pointed out that the value of expert opinions varied considerably among story types (for example, sports, international politics, and tech news). Lehmann et al. (2013a) showed that “news curators” who had expertise in particular topics played a key role, and distinguished them from those who tweeted about a lot of stories in less depth. In a related paper, Lehman et al. (2013b) suggested the “transient news crowds” that assembled around particular topics were the most useful reference group for journalists, reinforcing the idea of the fluid nature of news groups and networks on social media.

In an experimental test of which method has more predictive accuracy in spotting topics, Kang, O'Donovan, and Hollerer (2012) showed that the social model (using influencers) outperformed content-based predictions (such as keywords) and hybrid models. Using news about Libya during the uprising as a case study, they judged

that the social model had an accuracy score of 88.17%, significantly above the hybrid model (67%) and content-based model (63%).

Existing research supports the three key hypotheses that form the working assumptions in the Social Sensor project's approach to news and social media:

1. Twitter is the most important source for identifying news on social media.
2. Influencers have a key role in determining what news is likely to be most important in social media networks.
3. There is a need for a variety of methods, using both automated and manual techniques, to tackle the problems of verification.

Social Media and the News

Breaking news on social media

Over the last few years, social media has become a primary news source for journalists and ordinary news readers alike. Social networks are often the place where news is broken first and, increasingly, politicians, sports, and entertainment stars use social media channels to reach out to their followers but also to make announcements that they know will then be picked up by the media. As the size of social networks has grown, these channels have been increasingly used for distributing primary eyewitness material on breaking news stories. The Hudson River plane crash in 2009 was an early example of how a mobile phone picture can be distributed to a global audience within seconds—beating professional journalists to the story.

Since then, user-generated content posted to social networks like Twitter, Facebook, and YouTube has shaped coverage of news events, including the 2009 Iranian elections, the death of Osama Bin Laden in 2010, the Japanese earthquake of 2011, and the popular uprisings in the middle-east. As new technology has become widespread and audiences have become more familiar with these networks, the volume of user-generated content has grown exponentially.

During Hurricane Sandy in 2012, users of Instagram—a social photo sharing site—posted 10 photographs a second of the devastation, with around half a million photos being posted in total (Laird 2012). After the Boston Marathon bombings of 2013, Twitter carried hundreds of comments and messages of sympathy along with important official information. The capture of the second suspect was first announced in social media via the official police account (see figure 1). Social sources have become indispensable for the modern professional journalist. Most news organisations expect journalists to be fluent in social media in order to discover and distribute news, and many expect them to hold conversations with audiences as part of the ongoing production process.

[Insert figure 1]

Problems of Verification

The rise of social media has not been without its critics. Some point to the

unreliability of information in social networks and complain that the quality of information is being undermined by a growing 'cult of the amateur' and that rumours and falsehoods can be instantly spread around the world (Keen 2007). These issues have become pressing as mainstream media outlets integrate more and more social media content into their output and ordinary people increasingly use social media as a source of news. A 2013 survey of online news users in the UK showed that, on average, 25% used social media to find news at least once a week, but that less than 10% trusted that information (Reuters Institute 2013).

Over the years there have been a number of well-documented cases where misleading pictures and stories in social media have been given the 'oxygen of publicity' by news companies desperate to get one step ahead on a major news story. Following the death of Osama Bin Laden in May 2011, 'photoshopped' pictures purporting to show his dead body were distributed on social media and picked up by newspapers, news websites, and TV stations, potentially inflaming passions in the region (Newman 2011). Pictures are sometimes fabricated by governments or other official sources and then released via social media, on their own websites, or direct to news agencies. In 2008 AFP withdrew a picture of an Iranian missile test that had been changed to make the test look more impressive (Oliver 2008).

Alongside the deluge of genuine pictures of 2012's Hurricane Sandy, a large number of fake photographs (see figure 2) were also posted and a number of these were picked up by mainstream media outlets or retweeted by journalists (and thus given authenticity). It took painstaking analysis by several groups of journalists to confirm which images were faked (Madrigal 2012). An analysis of the top 100 most-tweeted picture stories for *The Guardian* datablog showed that 15% were fakes (Burgess, Vis, and Bruns 2012). In the December 2012 Connecticut school shootings, news agency and TV outlets misidentified the gunman to millions of readers and viewers, based on false information in social media (Zurawick 2012). Following the Boston Marathon bombings, social media tried to crowdsource the identifying of the perpetrators with unsatisfactory results. Reddit identified innocent people as suspects, and both Twitter and 4Chan were alive with both misinformation (Shih 2013) and outrage (Lee 2013) at the mistakes of both traditional and social media. The *New York Post*, CNN, and the AP were amongst the old media companies to have put out misleading information.

[Insert figure 2]

Pressure on Journalists

None of these issues are entirely new but they are becoming more complex in a world where the volume and speed of information are growing rapidly. Social media has raised a number of issues about how eyewitness material can be used and credited—but also about what kind of checks should be made to ensure veracity.

Our own research suggests that journalists have mixed feelings about social media. In early 2012 we interviewed 22 journalists from major news organisations such as the BBC, *The Guardian*, the *New York Times*, CNN and MSNBC, using in-depth, semi-structured interviews. These were supplemented by follow-up discussions with social media editors at a summit meeting in New York in April 2013. An analysis of

the language they used showed that journalists were convinced about the transformative effects of social media but often felt overwhelmed by the scale and speed of the changes (see figure 3).

[Insert figure 3]

These are a selection of comments made by journalists during the interviews:

It's really hard to find the nuggets of useful stuff in an ocean of content.

Twitter search is very hit and miss. What it produces is not comprehensive and the filters are not comprehensive enough.

Social media is huge and just being able to see the stuff that was relevant to us (a serious news organisation) and then further filter on the best people or comments would be ideal.

A real time search and discovery interface that just surfaced news and trends from multiple networks would just be brilliant.

Each journalist was asked to score, on a scale of 1-10, the relevance of a piece of social media functionality to them. Here are their preferences in order of importance:

1. Predicting or alerting breaking news **9.2**
2. Verifying social media content—quickly identifying who has posted a tweet or video and establishing “truth or lie” **8.5**
3. Listening—following high quality people / networks to find out about interesting / relevant stories **8.5**
4. Tracking trends and sentiment to inform programmes, news agenda **8.2**
5. Easily distribute your own content, find a new audience, and get feedback on what people thought of it **8.2**
6. Getting quick access to eyewitnesses or other trustworthy informants **7.55**
7. Crowdsourcing questions about a story in development—asking network for advice **5.2**

From this list it is clear that journalists use social media for a variety of day-to-day functions and this often relates to their role. While general journalists were relatively happy with the tools, social media specialists and senior editors were far from happy with the current level of technology support and workflow.

Existing tools for verification and the room for improvement

We also looked at a range of tools currently being used in newsrooms, and conducted a competitive analysis of the landscape. In 2012 we found that whilst ordinary journalists were using Twitter directly, editors and social media specialists were using more powerful tools such as Tweetdeck and Hootsuite, which allowed for multiple lists and filtering options. Some were also using specialist software such as Trendsmap and Google Trends to spot new stories. Tools for verification of contributors, such as Klout or PeerIndex, come up with aggregate scores for a contributor, but these were considered by our interviewees insufficiently granular to help journalists make judgments on authenticity in a fast-moving news story.

Most news organisations have published guidelines on how to manage these issues, and some have specific roles such as social media editor, or a group of people who are expert in managing this material, such as the BBC's UGC hub (Turner 2012) and the team behind CNN's iReport (Silverman 2012). Following the Costa Concordia disaster in 2011, the BBC social media team tried out different search terms on Twitter and looked for images on Twitpic and yfrog, and on YouTube for video. They tried to anticipate words people might use, such as 'sinking' and 'rocks'. They used geolocation services to drill down to people who might have been nearby (Trushar Barot, personal communication, January 2013). But currently, few posts are geo-tagged and tools often do not have enough fine-grained control. With the Norway bombings in July 2011, a CNN social media producer found an eyewitness who had uploaded a video to YouTube within a few minutes of the story breaking. They contacted and vetted him, and identified his Twitter location as part of the confirmation process. They then were able to use him on air speaking English and later speaking Spanish (various, personal communication and observation, January 2013).

The biggest news organisations continue to try to verify every picture or video they plan to use by contacting the owner directly. They do this to protect their editorial integrity but also because rights and payment for newsworthy footage are increasingly factors. By 2013, the volume of material and the speed with which they were able to verify it were becoming significant frustrations and, in most cases, smaller news organisations simply don't have the manpower to carry out these checks.

The Boston bombings have further focused attention on the need for verification and trust measures in social networks. Many senior editors have suggested that social networks themselves need to change the way they are organised in order to work more effectively with news organisations. Peter Horrocks, the BBC's Director of Global News, has called for a Moody's / Fitch style trust rating system for social media content (Hayward 2013a). Other editors are interested in labeling unverified material with hashtags, embedding journalists in social media networks to curate material, and producing a labeled trust and verification system, perhaps based on the eBay model (Hayward 2013b). Social media websites could also become self-policing. CCTV's Guo Chun reports that social media site Weibo has implemented punishments for people who are found to put out unreliable information, such as banning them from posting (Guo Chun, personal communication, April 20, 2013).

The Social Sensor Approach

For professional news journalists and general news users the key challenges around social media are the same: how to make sense of a huge amount of information and to identify relevant and important news quickly and with context. The aims of the social sensor project in relation to news are to:

1. Identify the key trends around social media in real time in a way that is useful and specific to public-interest journalism.
2. Identify key influencers and opinion-formers around any event and provide contextual information so journalists can make their own judgments.

3. Create a simple way to verify or authenticate user-generated content (text, images, video and audio) from social media sources.

Some Early Results – Finding Trending News in the US Election

One of our first experimental attempts to analyse the effectiveness of the tools Social Sensor is developing came with the US election campaign in 2012. We have concentrated on Twitter as our source of social media in our early experiments, as it is recognised as a key source by journalists and because its open API provides a ready source of data (see Aiello et al. [2013]).

To explore methods in crawling Twitter and topic detection, we wanted to create a substantial body of representative tweets, centred on a significant real-world event. We conducted two simultaneous crawls of Twitter on Election Day during the 2012 US Presidential elections. In one crawl, we collected all Tweets sent by (or to) a set of 5,000 “news hounds”. These were Twitter accounts that an algorithm had selected as being likely to tweet about US political issues. The accounts included American politicians and journalists, international journalists, and a wide range of other commentators, bloggers, and opinion-formers. In 12 hours during election night, we collected nearly 4 million tweets in this way. At the same time, in the other crawl, we collected nearly 2 million tweets that contained any of 20 hand-chosen keywords, which consisted primarily of candidates’ names and campaign hashtags. Examples included: #Election2012, #obama, #romney, #tcot, and #teaparty. At the peak, around 4:15am GMT (8:15 pm Pacific Standard Time in the US), we were collecting nearly 1,000 tweets per second in the news hounds’ crawl. One of the news hounds chosen by the algorithm was President Barack Obama. His victory tweet—“This happened because of you. Thank you”—was retweeted over 250,000 times very rapidly. In the five minutes immediately after that tweet was sent, our news hounds’ crawl collected the same message 51,910 times, accounting for over 28% of all tweets collected in that period. Those retweets dominated that period of our collection as they temporarily dominated Twitter (see figure 4).

[Insert figure 4]

Earlier in the evening, before polling stations closed, the campaign teams were still trying to get their supporters to vote. Retweets of Obama’s “If you’re in line when the polls close, stay in line to vote. It could help make the difference” made up over 10% of tweets collected over the next 15 minutes. Later, the official Obama account posted: “RT if you’re on #TeamObama tonight.” This was retweeted over 12,000 times and formed over 26% of the tweets we collected in the next hour (see figure 5). These results show the potential that key influencers have to dominate streams of messages, at least over short periods of time, as well as showing how campaigners can engage with their supporters directly.

[Insert figure 5]

While other contributors may not match Obama’s influence, they can still have an impact. When analyst Nate Silver (tweeting as @fivethirtyeight), the well-known *New York Times* columnist and data blogger, suggested earlier in the evening that

Obama's victory was now inevitable, his message, "On The Wall, The Writing", received over 4,000 retweets in just a few minutes (see figure 6).

[Insert figure 6]

Our software can find sudden bursts of activity by comparing the changes in the language used on Twitter from one minute to the next. These bursts of activity indicate newly-trending topics. Further technical details of this approach and a comparison of its performance with other state-of-art techniques can be found in Aiello et al. (2013). Earlier in the evening, our crawls collected tweets from news agencies and individual journalists who reported on the Electoral College results along with various senatorial races and referenda. Our software then automatically detected many of these as distinct stories. By focusing on selected individuals or specific keywords, we can largely exclude celebrity gossip and personal chatter, leaving predominantly newsworthy topics in near real-time.

Both methods of crawling—via news hounds or via a set of keywords—can produce a large number of newsworthy tweets. One advantage of the news hounds method is that we do not have to specify in advance what information we are seeking: we are effectively listening in on the conversations that thousands of experts are having with each of their audiences. We collect their audience's messages too, including links from eyewitnesses to related text, images and videos. Switching between different groups of news hounds allows us to rapidly change focus, for example between different countries or regions. As with any search engine, a keyword-based crawl can be very effective if suitable keywords can be chosen. For example, during the recent G8 summit we found many high-quality newsworthy tweets that contained the word "G8". But when an unexpected story is just emerging, it may be some time before such convenient search terms become apparent.

The earlier examples of Obama's tweets and consequent retweets were found in both the news hound and keyword crawls. However, Nate Silver's message was hardly found in the keyword collection because he did not name-check the president. Only when people retweeted it and appended hashtags such as #Obama2012 or #election2012 did it appear. In practice, a combination of news hounds and keywords seems to be most effective at finding useful content in real-time on Twitter.

Although the system of detecting trending news stories is still under development and prototype testing, these first results do suggest that identifying key influencers or news hounds who shape the tone of the coverage may be the most effective way of identifying trends. In some ways, this is an experimental confirmation of one of the fundamental theories of media research, the two-step model which predicts that mass opinion crucially depends on the role of trusted and influential individuals who are able to interpret and draw others' attention to key news developments (Lazarsfeld, Berelson, and Gaudet 1944; Katz and Lazarsfeld 1955).

Developing verification tools

'Alethia' is the Greek word for 'truth', and in Social Sensor, in conjunction with our partners at the Athens Technology Centre and Deutsche Welle, the German international public broadcaster, we are developing an 'Alethiometer': a module attempting to meter the credibility of information coming from any social source by

examining the three Cs—Contributors, Content, and Context. These seek to measure three key dimensions of credibility: the reliability of contributors, the nature of the content, and the context in which the information is presented. This reflects the range of considerations that working journalists take into account when trying to verify social media content. Each of these will be measured by multiple metrics based on our research into the steps that journalists go through manually. The results from the above steps can be weighted and combined to provide a sense of credibility to guide journalists. Our user research indicates that verification or guides to credibility are of crucial importance to journalists under severe time pressure. Further checks will often be necessary, but indications of credible content and ‘flags’ to highlight problematic content (as a way of preventing hoaxes) are likely to be of significant value.

We provide an initial set of five metrics for every one of the three categories which are being investigated for inclusion in the prototypes (figure 7 shows our prototype contributor verification page). Together with the metrics we provide an initial quantisation scheme for assigning credits to each metric. For every metric, we intend to normalise the results on a 10-scale grading scheme: a negative result in any metric will raise a suspicious flag to be shown to the user. When the user clicks on a flag, the system will display the corresponding metric screen that raised the particular flag. These additional screens will give more detail and context about the Contributor, the Content and the Context. Over time we will be able to build up information based on multiple publicly-available sources to help journalists understand the context of a tweet or picture. For example, showing the geographic location of the followers of a contributor can give clues about their likely location, which may affect their credibility on a particular story. The next stage of our research is to test the concept of the Alethiometer with working journalists, using an end-to-end system. The evaluation is done through a variety of metrics, including internal and external rankings and comparison with other data sources.

[Insert figure 7]

Conclusions

The problems of identifying key trends and verifying social media news have challenged media organisations and software designers alike. Collaboration between social scientists, journalists, and computer scientists, as in the Social Sensor project, can make a significant difference to the way new tools are conceptualised, developed and tested. By relying on the needs and practices of working journalists, and using a user-centric and iterative process of prototype development, the project hopes to make a significant contribution to addressing these problems by developing innovative and flexible tools and usable heuristic metrics.

Acknowledgments

Our thanks to the European Union for supporting this research as part of the Social Sensor consortium FP7 research grant 287975.

References

Aiello, Luca Maria, G. Petkos, C. Martin, D. Corney, S. Papadopoulos, R. Skraba, A.

Goker, I. Kompatsiaris, and A. Jaimes. 2013. "Sensing Trending Topics in Twitter." *IEEE Transactions on Multimedia*, Issue 99, <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6525357&isnumber=4456689>

Burgess, Jean, Farida Vis, and Axel Bruns. 2012. "Hurricane Sandy: the Most Tweeted Pictures." *The Guardian Data Blog*, November 6. <http://www.guardian.co.uk/news/datablog/gallery/2012/nov/06/hurricane-sandy-tweeted-pictures>

Castillo, Carlos, Mohammed El-Haddad, Jurgen Pfeffer, and Matt Stempeck. 2012. "Characterizing the Life Cycle of Online News Stories Using Social Media Reactions." [unpublished]

Castillo, Carlos, Marcelo Mendoza, and Barbara Poblete. In press. "Predicting Information Credibility in Time-Sensitive Social Media." *Internet Research, special issue on The Predictive Power of Social Media*.

Diakopoulos, Nicholas, Munmun De Choudhury, and Mor Naaman. 2012. "Finding and Assessing Social Media Information Sources in the Context of Journalism." *Proceedings, CHI'12, Austin, Texas, May*.

Hahn, Nadja. 2013. "What Good is Twitter? The Value of Social Media to Public Service Journalists." *Polis*, March 2013.

Hayward, David. 2013a. "After Boston, Brainstorming Better Verification #smsnyc." *BBC Academy*, April 22. <http://www.bbc.co.uk/blogs/blogcollegeofjournalism/posts/After-Boston-brainstorming-better-verification-smsnyc>

Hayward, David. 2013b. "Future of News Organisations in a Social World: The Big Guns Discuss #smsnyc." *BBC Academy*, April 23. <http://www.bbc.co.uk/blogs/blogcollegeofjournalism/posts/Future-of-news-organisations-in-a-social-world-the-big-guns-discuss-smsnyc>

Hsieh, Chu-Cheng, Christopher Moghbel, Jianhong Fang, Junghoo Cho. 2013. "Experts vs The Crowd: Examining Popular News Prediction Performance on Twitter." *Proceedings of the WWW13 conference, Rio de Janeiro, May*.

Kang, Byungkyu, John O'Donovan, and Tobias Hollerer. 2012. "Modeling Topic Specific Credibility on Twitter." *Proceedings IUI'12, Lisbon, February*.

Katz, E., and P. Lazarsfeld. 1955. *Personal Influence*. New York: The Free Press.

Kwak, Haewoon, Changhyun Lee, Hosung Park, and Sue Moon. 2010. "What is Twitter, a Social Network or a News Media?" *Proceedings, WWW10, Raleigh, North Carolina, April*.

Keen, Andrew. 2007. *The Cult of the Amateur: How Today's Internet is Killing Our Culture*. London: Broadway Business.

- Laird, Sam. 2012. "Instagram Users Share 10 Hurricane Sandy Photos Per Second." *Mashable*, October 29. <http://mashable.com/2012/10/29/instagram-hurricane-sandy/>
- Lasorsa, Dominic, Seth Lewis, and Avery Holton. 2012. "Normalising Twitter." *Journalism Studies* 13:1, 19-36.
- Lazarsfeld, P.F., B. Berelson, and H. Gaudet. 1944. *The People's Choice: How the Voter Makes up his Mind in a Presidential Campaign*. New York: Columbia University Press.
- Lee, Dave. 2013. "Boston Bombing: How Internet Detectives Got it Very Wrong." *BBC News*, April 19. <http://www.bbc.co.uk/news/technology-22214511>
- Lehmann, Janette, Carlos Castillo, Mounia Lalmas, and Ethan Zuckerman. 2013a. "Finding News Curators in Twitter." *Proceedings of the WWW13 conference, Rio de Janeiro, May (Social News on the Web workshop)*.
- Lehmann, Janette, Carlos Castillo, Mounia Lalmas and Ethan Zuckerman. 2013b. "Transient News Crowds in Social Media." *Proceedings of the WWW13 conference, Rio de Janeiro, May*.
- Madrigal, Alexis. 2012. "Sorting the Real Sandy Photos from the Fakes." *The Atlantic*, October 29. <http://www.theatlantic.com/technology/archive/2012/10/sorting-the-real-sandy-photos-from-the-fakes/264243/>
- Morales, Gianmarco De Francisci, Aristides Gionis, and Claudio Lucchese. 2012. "From Chatter to Headlines: Harnessing the Real-Time Web for Personalized News Recommendations." *Proceedings WSDM'12, Seattle. Washington, February*
- Newman, Nic. 2009. "The Rise of Social Media and its Impact on Mainstream Journalism." Working Paper. Reuters Institute for the Study of Journalism, Oxford University.
- Newman, Nic. 2011. "The Mainstream Media and the Distribution of News in the Age of Social Discovery." Working paper. Reuters Institute for the Study of Journalism, Oxford University.
- Newman, Nic, William Dutton, and Grant Blank. 2012. "Social Media in the Changing Ecology of News: The Fourth and Fifth Estates in Britain." *International Journal of Internet Science*, Vol 7:1.
- Oliver, Laura. 2008. "Spot the Difference: AFP Withdraws 'Digitally Altered' Missile Shot." *Journalism.co.uk*, July 10. <http://blogs.journalism.co.uk/2008/07/10/spot-the-difference-afp-withdraws-digitally-altered-missile-shot/>
- Petrovic, Sasa, Miles Osborne, Richard McCreddie, Craig Macdonald, Iadh Ounis, and Luke Shrimpton. 2013. "Can Twitter Replace Newswire for Breaking News?" *Association for the Advancement of Artificial Intelligence*

Pew Research Centre, Project for Excellence in Journalism. 2012. "What Facebook and Twitter Mean for News." *The State of the News Media Annual Report*, Washington DC

Reuters Institute for the Study of Journalism. 2013. "Digital News Report 2013." <http://digitalnewsreport.org>

Shih, Gerry. 2013. "Boston Marathon Bombings: How Twitter and Reddit Got it Wrong." *The Independent*, April 20.
<http://www.independent.co.uk/news/world/americas/boston-marathon-bombings-how-twitter-and-reddit-got-it-wrong-8581167.html>

Silverman, Craig. 2012. "How CNN's iReport verifies its Citizen Content." *Poynter*, January 26, <http://www.poynter.org/latest-news/regret-the-error/160045/how-cnns-ireport-verifies-its-citizen-content/>

Turner, Dave. 2012. "Inside the BBC's Verification Hub." *Nieman Reports (summer 2012)* <http://www.nieman.harvard.edu/reports/article/102764/Inside-the-BBCs-Verification-Hub.aspx>

Zurawick, David. 2012. "School Shooting: A Confused Media Fail Us Again." *Daily Download*, December 15. <http://daily-download.com/school-shooting-confused-media-fail/>

Figures

Figure 1: Twitter report from Boston Police Department on capture of Boston Marathon bomber.



Figure 2: Example of ‘photoshopped’ fake photograph that circulated on social media during 2012’s Hurricane Sandy

Like I said.... I wish I were there. I should have been a storm chaser.

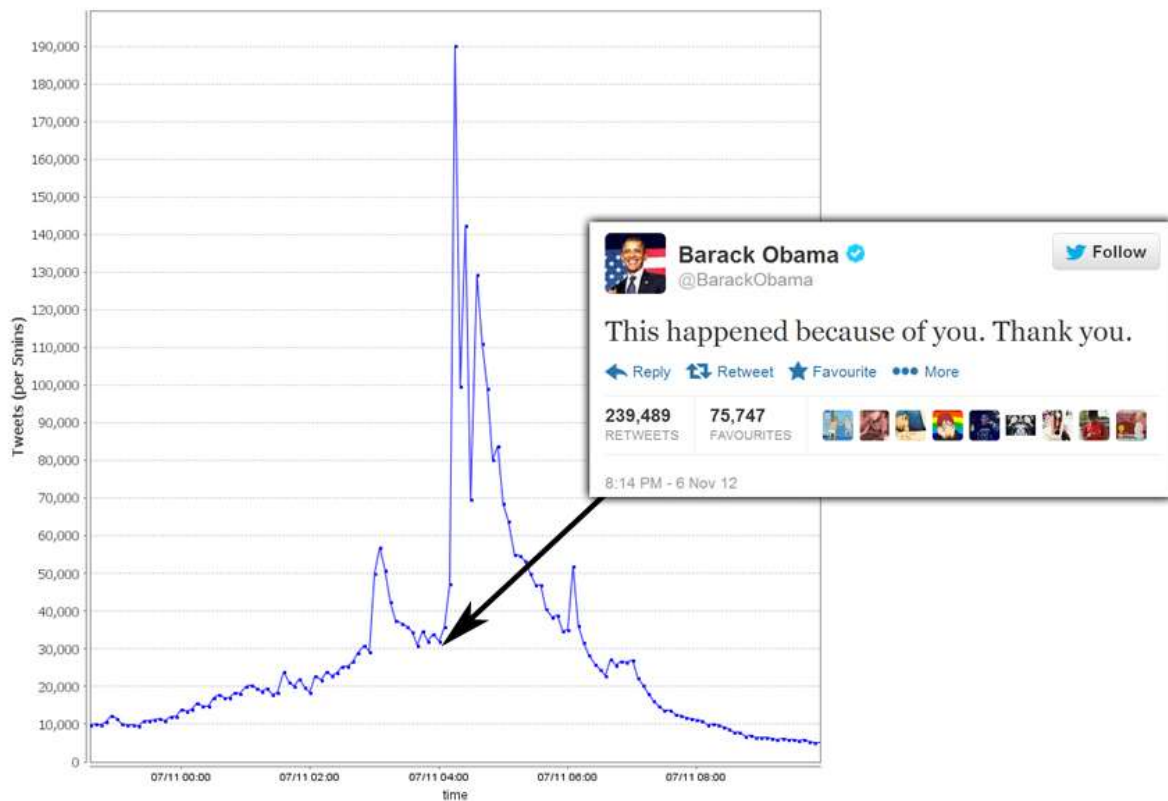


This is an amazing shot of New York today with the Frankenstorm bearing down. Nature is so powerful, yet so beautiful.

Figure 3: The most common words used by journalists interviewed on the impact of social media on their work

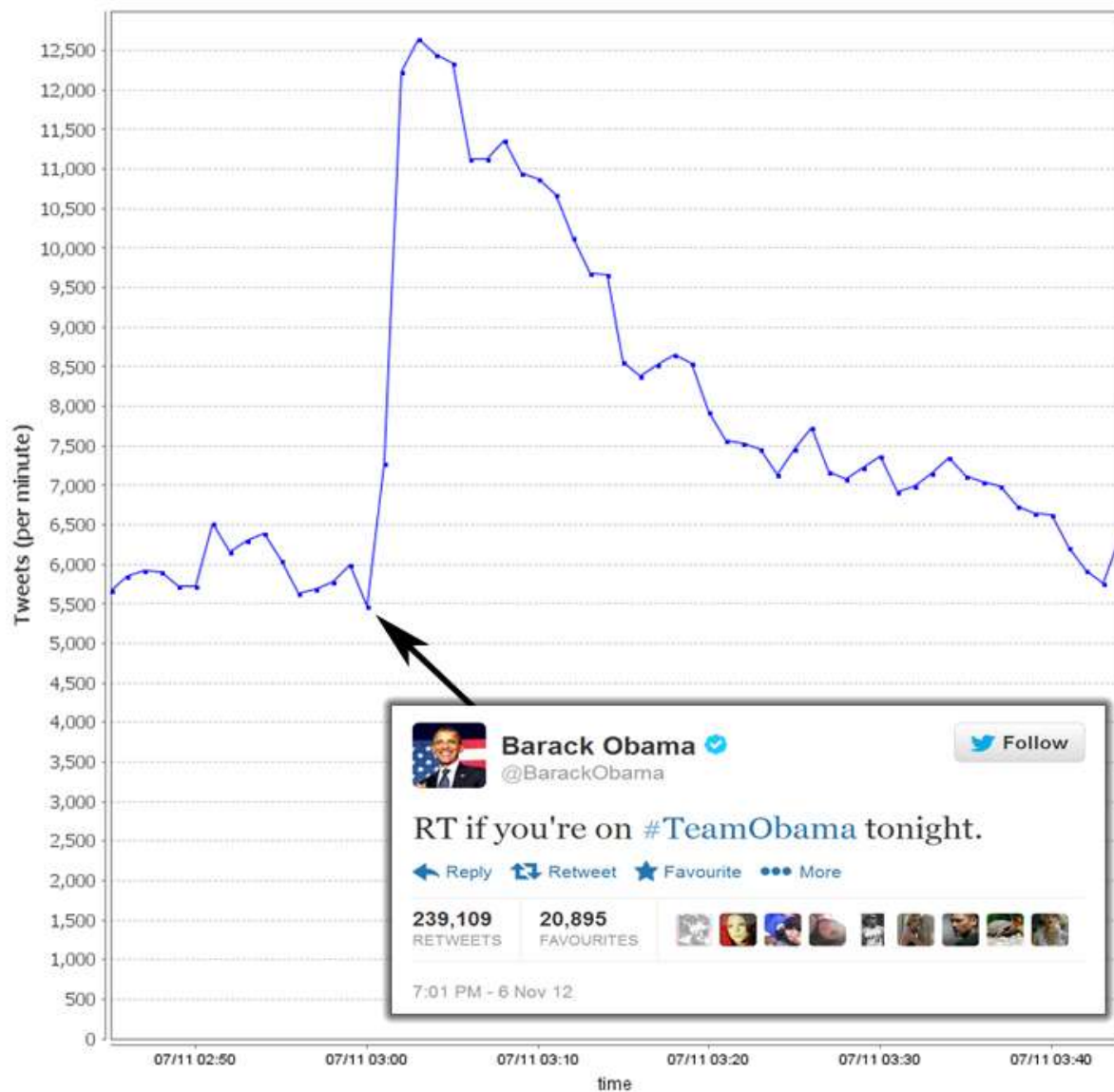


Figure 4: Number of retweets every five minutes of Barack Obama's victory tweet—"This happened because of you. Thank you"—on US election night, November 6, 2012.



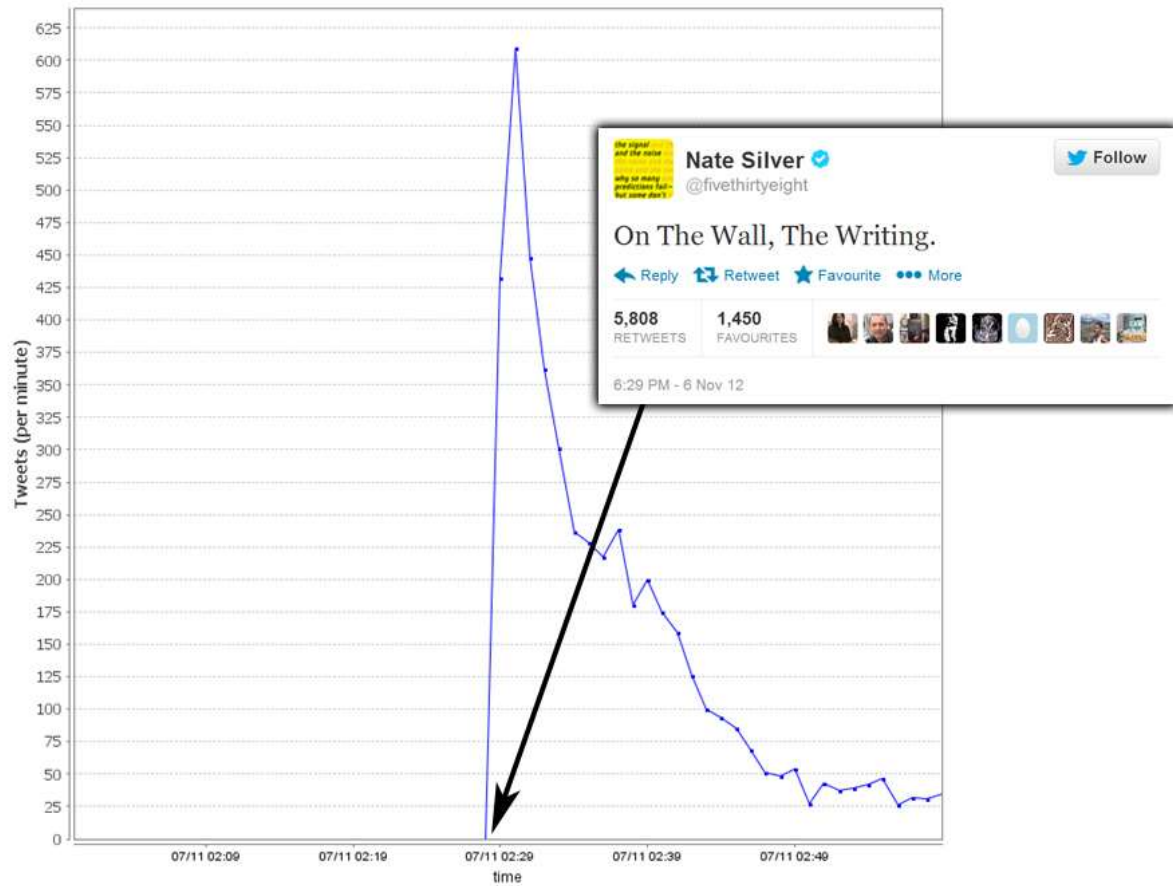
Note: X-axis is GMT. Time of Tweet shown in US Pacific Standard Time.

Figure 5: Number of retweets every minute of Barack Obama’s tweet—“RT if you’re on #TeamObama tonight”—on US election night, November 6, 2012.



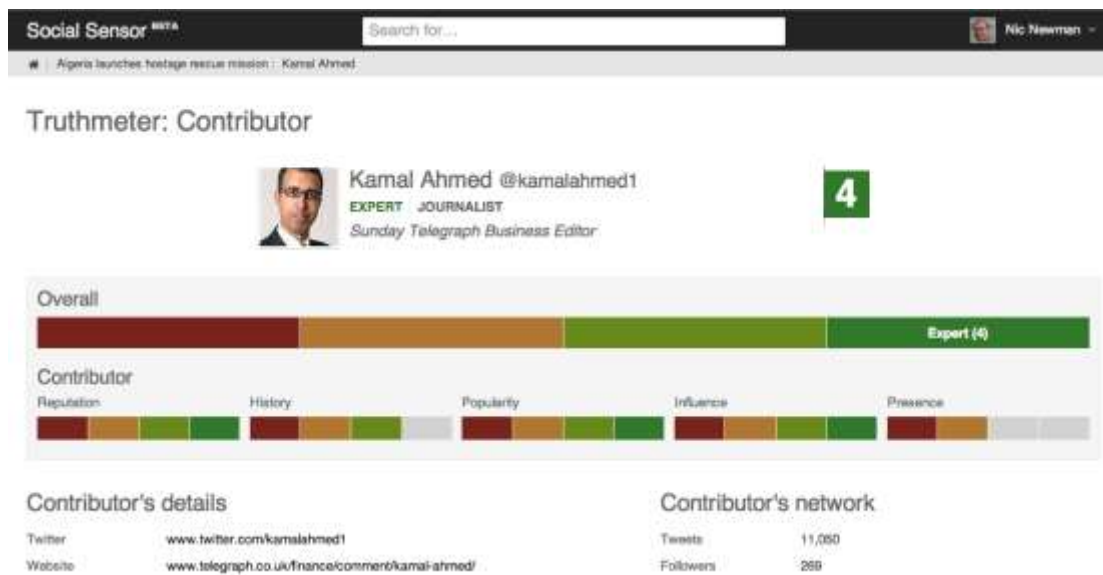
Note: X-axis is GMT. Time of Tweet shown in US Pacific Standard Time.

Figure 6: Number of retweets every minute of Nate Silver's tweet—"On The Wall, The Writing"—on US election night, November 6, 2012.



Note: X-axis is GMT. Time of Tweet shown in US Pacific Standard Time.

Figure 7: The Social Sensor news tool prototype: contributor verification page



¹ For further details on the Social Sensor research consortium see <http://www.socialsensor.eu>