



Archived Content

In an effort to keep DHS.gov current, this document has been archived and contains outdated information that may not reflect current policy or programs.



Integrated Public Alert and Warning System

Wireless Emergency Alerts

Understand and Respond to Public Sentiment

First Responders Group
November 30, 2014



**Homeland
Security**

Science and Technology



Integrated Public Alert and Warning System
Wireless Emergency Alerts
Understand and Respond to Public Sentiment

HS HSSEDI™ Task
HSHQDC-13-J-00097

Version 4.0

This document is a product of the Homeland Security Systems Engineering and Development Institute (HS HSSEDI™).

ACKNOWLEDGEMENTS

The Homeland Security Systems Engineering and Development Institute (hereafter “HS HSEDI” or “HSEDI”) is a federally funded research and development center established by the Secretary of Homeland Security under Section 305 of the Homeland Security Act of 2002. The MITRE Corporation operates HSEDI under the Department of Homeland Security (DHS) contract number HSHQDC- 09-D-00001.

HSEDI’s mission is to assist the Secretary of Homeland Security, the Under Secretary for Science and Technology, and the DHS operating elements in addressing national homeland security system development issues where technical and systems engineering expertise is required. HSEDI also consults with other government agencies, nongovernmental organizations, institutions of higher education and nonprofit organizations. HSEDI delivers independent and objective analyses and advice to support systems development, decision making, alternative approaches and new insight into significant acquisition issues. HSEDI’s research is undertaken by mutual consent with DHS and is organized by tasks in the annual HSEDI Research Plan.

This report presents the results of test planning conducted under **HSHQDC-13-J-00097, Science and Technology Directorate (S&T) FEMA Integrated Public Alert and Warning System (IPAWS)/Wireless Emergency Alert (WEA) Test and Evaluation** of HSEDI’s Fiscal Year 2013 Research Plan. The purpose of the task is to apply independent and objective system engineering expertise and provide technical advice, guidance and subject matter expertise in the performance of test planning and execution to verify the WEA technical requirements.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of DHS.

For more information about this publication contact:

Homeland Security Systems Engineering and Development Institute

The MITRE Corporation

7515 Colshire Drive

McLean, VA 22102

Email: hssedi_info@mitre.org

www.mitre.org/HSEDI

Abstract

The Integrated Public Alert and Warning System (IPAWS) is the Federal Emergency Management Agency's system used to distribute alert and warning messages. The IPAWS system provides public safety officials with a capability to alert the public about emergencies using the Emergency Alert System, Wireless Emergency Alerts (WEA), the National Weather Service (NWS) Weather Radio, and other public alerting services.

As there is now a growing number of people who receive WEA messages through their commercial service providers and on their cellular phones, it is an opportune time to assess the public's sentiment towards the WEA messages they receive and help improve the communication of information. Analyzing sentiment can be done by using industry tools to gain insight about how the public uses online and social media sources to discuss and react to the alerts. The Homeland Security Systems Engineering and Development Institute (HSSEDI) executed a proof-of-concept to demonstrate how to analyze public sentiment vis-à-vis WEA messages using commercial and HSSEDI research tools. Commercial tools were used to access social media and to perform basic analytics. These tools allowed selected data to be exported for further analysis in the HSSEDI research tools. The purpose of the analysis was to understand the expressed sentiment across online and social media in reaction to WEA messages and to make recommendations to Alert Originators on techniques to handle negative sentiment in reaction to WEA messages.

HSSEDI also observed how NWS sends weather messages and made recommendations on how to improve public interactions.

Key Words

- Advanced Analytics
- Alert Originator
- America's Missing: Broadcast Emergency Response alerts
- Integrated Public Alert and Warning
- National Oceanographic and Atmosphere Administration
- National Weather Service
- Online and Social Media Sources
- Sentiment Analysis
- Simulation Models
- Wireless Emergency Alerts

Record of Changes

| No. | Date | Reference | A=Add M=Modify D=Delete | Description of Change |
|-----|--------------|--|-------------------------------|--|
| 1 | Oct 15, 2014 | Jennifer Mathieu Ryan Hollins Holly Babin Ellen Greene John James | A | Formal Deliverable: IPAWS, WEA: Understand and Respond to Public Perception |
| 2 | Oct 28, 2014 | Jennifer Mathieu Ryan Hollins Allison Ounanian Holly Babin Ellen Greene John James Paula Mahoney Gary Fleming Donna Trammell | M | Informal Deliverable: Updated IPAWS, WEA: Understand and Respond to Public Perception to include detailed information |
| 3 | Nov 14, 2014 | Jennifer Mathieu Ryan Hollins Allison Ounanian Ellen Greene John James Paula Mahoney Donna Trammell Ben Schimmel Brian Wilkins John Dewsnap Allen Boetig | M | Formal Deliverable: Updated IPAWS, WEA: Understand and Respond to Public Perception to include Crimson Hexagon analysis, Sentimdir results, and model results, as well as addressed both sponsor and peer reviewed comments |
| 4 | Nov 26, 2014 | Jennifer Mathieu Ryan Hollins Allison Ounanian Ellen Greene John James Paula Mahoney Ben Schimmel Brian Wilkins John Dewsnap Allen Boetig Don McGarry | M | Informal Deliverable: Updated IPAWS, WEA: Understand and Respond to Public Perception to include New NWS/NWS section; new model results, and new industry tool evaluation summary, as well as addressed both sponsor and peer reviewed comments |
| 5 | Nov 30, 2014 | All | M | Final Update. Clarified various sections and rearranged appendixes. |

1 Executive Summary

The Wireless Emergency Alerts (WEA) service has been in use since April 2012, and there is a growing base of commercial mobile service subscribers who receive WEA messages. Therefore, it is an opportune time to assess to the public's sentiment towards WEA and improve the communication of information related to the alert. Public sentiment towards WEA messages can be assessed by analyzing user discussions in online (e.g., news, blogs) and social media (e.g., Twitter, Facebook) platforms.

To gauge public sentiment towards WEA messages, the Homeland Security Systems Engineering and Development Institute (HSSEDI) used commercial tools to search online media sources for user comments and feedback. MITRE-developed research tools then analyzed this information to gain an understanding of the public sentiment related to WEA messages.

HSSEDI executed a proof-of-concept to demonstrate how to analyze public sentiment towards WEA messages using commercial and HSSEDI research tools. Commercial tools were used to access social media and perform basic analytics. These tools allowed selected data to be exported for further analysis in the HSSEDI research tools. The analysis sought to understand the expressed sentiment across online and social media in reaction to WEA messages and to make recommendations to alert originators on techniques to handle negative sentiment.

HSSEDI also observed how the National Weather Service (NWS) sends weather messages and made recommendations to improve public interactions.

Finally, HSSEDI performed a case study on the use of social media tools in real time to monitor social media activity. This case study was performed at one of the NWS field offices. The case study showed that for this facility, NWS personnel have limited time to monitor online and social media data sources, even if tools can be used to rapidly facilitate understanding public sentiment towards specific warnings. HSSEDI proposes leveraging existing trained weather spotters to assist in understanding and responding to public sentiment in real time.

Table of Contents

| | |
|--|-----------|
| Executive Summary | 6 |
| 1 Introduction..... | 10 |
| 1.1 IPAWS | 10 |
| 1.2 Social Media | 11 |
| 1.3 Tools | 12 |
| 2 Technical Approach..... | 13 |
| 2.1 Broad WEA Data Set Selection | 13 |
| 2.2 Social Media Mining Methodology | 13 |
| 2.3 Content Selection | 13 |
| 2.4 Content Analysis | 14 |
| 3 Analysis Summary | 15 |
| 3.1 General Observations..... | 15 |
| 3.2 Advanced Analytics | 16 |
| 4 Observation at NWS | 18 |
| 4.1 Overall Information | 18 |
| 4.2 Specific Recommendations for NWS | 18 |
| 5 Recommendations | 19 |
| 5.1 General Recommendations for WEA | 19 |
| 5.2 Engagement Methodology | 20 |
| 6 Conclusion | 23 |
| Appendix A | |
| Historical Analysis..... | 24 |
| Appendix B | |
| Tools | 34 |
| Appendix C | |
| ResearchTools | 49 |
| Appendix D | |
| National Weather Service Site Visit | 59 |
| Appendix | |
| E Acronyms | 65 |

List of Tables

Table 1: Current Research Tools 49

List of Figures

| | |
|--|----|
| Figure 1: IPAWS WEA System..... | 11 |
| Figure 2: Online and Social Media Landscape | 12 |
| Figure 3: Commercial and HSSEDI Research Tool Data Flow | 14 |
| Figure 4: Threshold in Distance for Negative Public Sentiment | 17 |
| Figure 5: Initial Engagement Methodology | 21 |
| Figure 6: A Tweet Expresses Annoyance at Large Number of Alerts | 25 |
| Figure 7: Grouped Ideas or Memes | 26 |
| Figure 8: Geographic Distribution of Tweet Postings | 26 |
| Figure 9: Two Peaks in Early and Mid-October | 27 |
| Figure 10: CoFi Analysis of First Peak Yields Kentucky Conversation | 28 |
| Figure 11: CoFi Analysis of Second Peak Yields Texas and Colorado Conversations..... | 28 |
| Figure 12: Multiple Breakpoints Over Two Days | 29 |
| Figure 13: CoFi Analysis of Hawaiian Flash Flood | 30 |
| Figure 14: Six AMBER Alerts with a Peak on January 11, 2014..... | 30 |
| Figure 15: CoFi Analysis Shows Alize Whipple's AMBER Alerts Dominates Conversation | 31 |
| Figure 16: Chatter Increases Quickly then Decreases as Case Status Changed | 32 |
| Figure 17: Spike in positive sentiment expressed in messages..... | 33 |
| Figure 18: Drilldown Using Sentimedir | 33 |
| Figure 20: Sentiment Over Time | 50 |
| Figure 21: Breakpoint Analysis with Tweet Explorer | 51 |
| Figure 22: CoFi Display..... | 53 |
| Figure 23: Pinocchio Display of Tweets with Similar Content | 54 |
| Figure 24: Event Detector Display Showing Detected Events from Tweets..... | 55 |
| Figure 25: NWS Radar with Alert Box..... | 59 |
| Figure 26: NWS Warning Software and Message Output..... | 60 |
| Figure 27: NWS Weather Message..... | 61 |
| Figure 28: Example Use of HootSuite to Follow Facebook and Twitter | 63 |

1 Introduction

The Homeland Security Systems Engineering and Development Institute (HSSEDI) performed an analysis to demonstrate how to analyze public sentiment towards Wireless Emergency Alerts (WEA) messages using commercial and HSSEDI research tools. Commercial tools were used to access online and social media and perform basic analytics. These tools allowed selected data to be exported for further analysis in the HSSEDI research tools. The analysis sought to understand the expressed sentiment across social media in reaction to WEA messages, and to make recommendations to Alert Originators (AOs) on techniques to respond to current negative sentiment and reduce future occurrences of negative sentiment. HSSEDI observed how National Weather Service (NWS) sends weather messages and made recommendations to improve public interactions.

1.1 IPAWS

The Integrated Public Alert and Warning System (IPAWS) is the Federal Emergency Management Agency's (FEMA) system used to distribute warning messages. The system uses the Emergency Alert System (EAS), WEA, the NWS Weather Radio, and other services to distribute the alerts.

IPAWS is a Federated¹ Family-of-Systems² (FFOS), an umbrella architecture and program for the next generation EAS that enables several pathways for alert distribution, including EAS (radio and TV) and WEA messages (cellular phones). Within the IPAWS FFOS, the FEMA-developed IPAWS Open Platform for Emergency Networks (IPAWS-OPEN) Alerts Aggregator is the central software component for distributing alerts generated by AOs.

As shown in Figure 1, IPAWS-OPEN enables the interoperable sharing of emergency alerts and incident-related data between various IPAWS alert and warning subsystems (including WEA) by collecting and routing IPAWS emergency alerts to and from the emergency systems that serve the public. For WEA, IPAWS-OPEN allows an AO to submit messages for distribution to the gateways of participating Commercial Mobile Service Providers (CMSP). In turn, the CMSPs disseminate WEA messages in a text message format to Commercial Mobile Service (CMS) Subscribers. Cell Broadcast (CB) technology³ is used to send the WEA messages via a different channel than normal text messages to ensure they will not get backlogged during times of emergency, when wireless voice and data services are highly congested. Furthermore, as an emergency situation unfolds, an AO also has the capability to generate updates and cancel emergency alerts for dissemination by the CMSPs.

¹ A federated architecture approach recognizes the uniqueness and specific purpose of disparate architectures and allows for their autonomy and local governance while enabling the enterprise to benefit from their content. "DoD Architecture Framework, Version 1.5." Department of Defense, 2007. Web. http://dodcio.defense.gov/Portals/0/Documents/DODAF/DoDAF_Volume_1.pdf.

² Family of Systems (FoS) - A set or arrangement of independent systems that can be arranged or interconnected in various ways to provide different capabilities. The mix of systems can be tailored to provide desired capabilities, dependent on the situation. "Family of Systems." *Defense Acquisition University (DAU)*. 21 Apr. 2005. Web. <https://acc.dau.mil/CommunityBrowser.aspx?id=54714>.

³ Cell Broadcast technology is a mechanism to broadcast a message to all CB-capable devices connected to a tower using an out-of-band communication mechanism. This allows the message to be broadcast without conflict with voice or data transmissions.

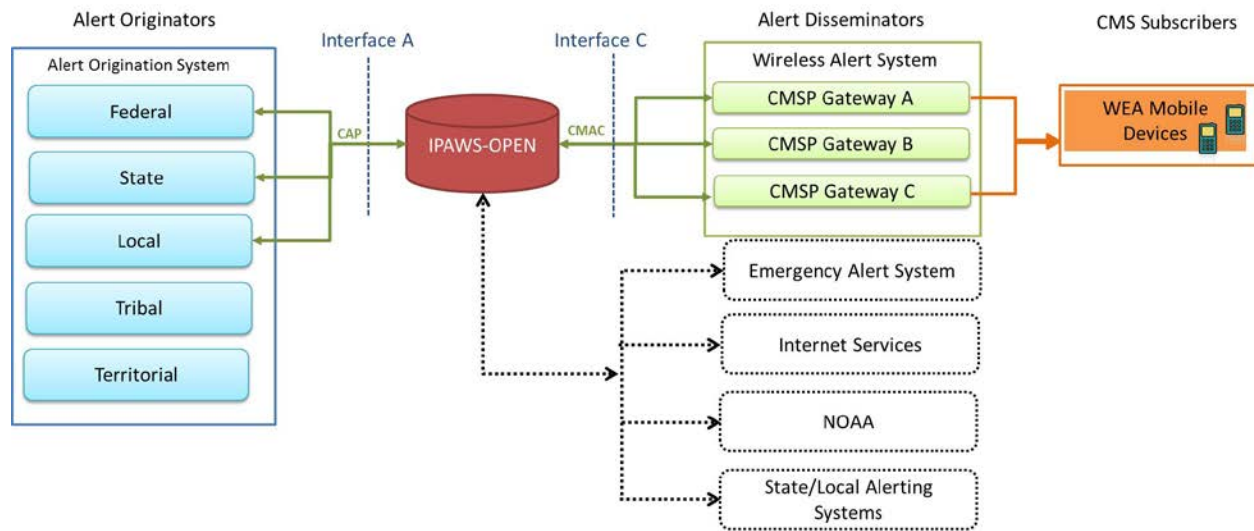


Figure 1: IPAWS WEA System

There are three categories of WEA messages: Presidential, Imminent Threat, and America's Missing: Broadcast Emergency Response (AMBER). A Presidential alert is used to warn or inform the American people of situations of national importance and is only sent by the President of the United States. An Imminent Threat alert is issued when there is a potential or expected loss of life, a potential for severe injury, or significant damage is likely or expected within 24 hours. AMBER alerts are sent to the public when there is a suspected abduction of a child and issued by National Center for Missing and Exploited Children (NCMEC).⁴

1.2 Social Media

The analysis of sentiment in tweets pertaining to the selected WEA messages was based on a set of 396 geo-located tweets (latitude and longitude coordinates from the users' device) between March 2013 and September 2014. The geo-located Twitter data was compared to weather alerts over the same time period. Figure 2 shows a sample of online and social media platform. For this analysis, Twitter was the primary source of information. Twitter is a social media platform that allows users to send and receive 140 character messages, which are called 'tweets.' Twitter was chosen for this analysis due to the ease of access to historical data.

⁴ <http://www.missingkids.com>



Figure 2: Online and Social Media Landscape

1.3 Tools

Since the amount of data available from any social media source is vast, data analysis tools were required to filter this information into usable amounts. The following commercial tools were used to analyze past content about WEA messages from Twitter and perform basic analysis:

- Crimson Hexagon
- Sysomos
- Gnip

Once a set of tweets was retrieved, research tools were required to conduct advanced analysis of the content. HSSDI used the following MITRE tools to perform advanced analysis on Twitter content:

- Sentimdir
- Tweet Explorer
- Comment Filter (CoFi)
- MemeME
- Pinocchio
- Timeline

Select analysis using the tools is shown in Appendix A and additional information on these tools can be found in Appendices B and C.

2 Technical Approach

To gauge public sentiment towards WEA messages, HSSEDI used commercial tools to search online media sources. MITRE-developed research tools then analyzed this information to gain an understanding of the public sentiment related to WEA messages.

2.1 Broad WEA Data Set Selection

The FEMA IPAWS Program Management Office (PMO) provided a historic log of 10,000 WEA messages for use in this analysis. The 10,000 alerts were plotted over time periods when WEA message volume was relatively high. The alerts in these timeframes were clustered⁵ by state and county. This clustering identified locations that had a large number of WEA messages. These locations were used in the analysis. Based on the review of the entire historic IPAWS log, HSSEDI observed that the vast majority of alerts issued were NWS⁶ Imminent Threat alerts⁷ with a small number of AMBER alerts and no Presidential alerts.

2.2 Social Media Mining Methodology

Social media content pertaining to public sentiment towards Imminent Threat or AMBER alerts was narrowed to Twitter, then filtered to minimize the number of tweets unrelated to these alerts. Twitter was queried for relevant conversations about the WEA messages occurring in the identified peak time periods. The queries contained keywords related to Imminent Threat weather alerts and AMBER alerts. Additionally, tweet geo-location was used to narrow down the resulting data set to the United States. Through several iterations of query execution, the queries were refined based on the results returned until usable sets of information were obtained.

2.3 Content Selection

Imminent Threat and AMBER alerts that generated more significant online and social media data were prioritized over those with little data available. Based on social media data available, the following Imminent Threat alerts were prioritized:

- Flash Floods: Alert effectiveness is very challenging with flash floods, since they can cover a vast geographic area, occur very quickly and result in substantial damage. For these reasons, flooding events were given higher priority.
- Blizzards: Due to the effectiveness of traditional news/weather stations in forecasting blizzards, this type of weather event takes a lower priority and is in fact no longer disseminated as WEA messages.
- Other weather events (e.g., tornados) were given standard priority.

⁵ "A group of similar things or people positioned or occurring closely together."
<http://dictionary.reference.com/browse/cluster>

⁶ Weather alerts are issued by NWS which is part of the National Oceanic and Atmospheric Administration (NOAA).

⁷ NOAA issues only weather alerts.

2.4 Content Analysis

Both commercial and HSSEDI research tools were used to perform analysis on the data. The figure below shows the dataflow between tools. In this figure, tan boxes identify commercial tools and blue boxes distinguish HSSEDI research tools. As shown in the figure, data exported from the commercial tools was used by the HSSEDI research tools.

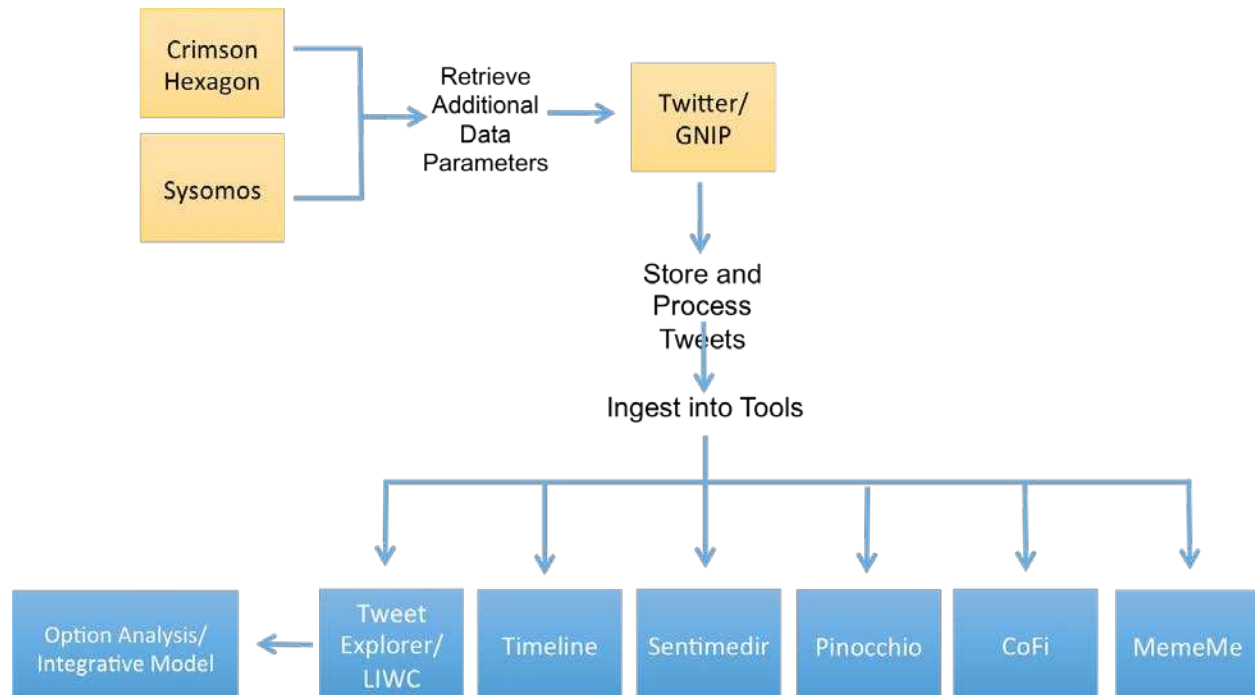


Figure 3: Commercial and HSSEDI Research Tool Data Flow

The HSSEDI research tools provided advanced analytics for the determination of public sentiment and reactions towards WEA messages. As part of the analysis, HSSEDI geo-located⁸ tweets and grouped them by distance from the center of the alert. The sentiment of the alert was then analyzed by this distance.

⁸ "Geo-location is the identification of the real-world geo-graphic location of an object, such as a radar source, mobile phone or Internet-connected computer terminal. Geo-location may refer to the practice of assessing the location, or to the actual assessed location." <http://en.wikipedia.org/wiki/Geolocation> (Nov 30, 2014).

3 Analysis Summary

Based on the prioritization method described in Section 2, HSSEDI chose the following alerts from the time period of October 2013 to April 2014 for analysis. Four data sets were selected for further study:

1. Tornadoes and flooding in Arkansas and Mississippi (April 27 – 30, 2014)
2. Severe weather in Texas and elsewhere (October 13 – 18, 2013)
3. Flash floods in Hawaii (December 30 – 31, 2013)
4. Six AMBER alerts covering multiple states (January 1 – 17, 2014)⁹
 - a) AMBER alert for Jagger Curry in Texas
 - b) AMBER alert for Alize Whipple in Massachusetts
 - c) AMBER alert for Daniel Britt in Rhode Island and Washington
 - d) AMBER alert for Elizabeth Smith in Texas
 - e) AMBER alert for Tristin Radey in Missouri
 - f) AMBER alert for Mitchell Farris in Missouri

HSSEDI's analysis suggested that AOs can address negative public sentiment towards WEA messages by educating the public about the WEA program and engaging to clarify inaccurate or incomplete information. The analysis showed a threshold, between 40 and 80 miles from the center of the alert area, where public sentiment becomes much more negative for a WEA message.

3.1 General Observations

The following sections contain observations and results of an analysis of the social media content of the Twitter data for Imminent Threat alerts and AMBER alerts. Broadly, these fall into categories: alert delivery times, audio tones and the relevance of content to recipients' locations. Both types of alerts (Imminent Threat and AMBER) had some negative sentiment expressed towards them, as well as some uncertainty surrounding the WEA messaging capability.

3.1.1 Imminent Threat Alerts

The following list contains the results of the analysis of online and social media sources for Imminent Threat alerts using commercial and HSSEDI research tools.

- Twitter users complained about the sound their cell phones made when the WEA messages were received.
- Twitter users said they were unaware their phones had such an alerting capability, or were unaware of how the alert was transmitted to their phone.
- Twitter users complained about the number of alerts they received.
- Twitter users complained about the times at which they received the alert; this was often during early morning hours.

⁹ These alerts were treated as one data set because the online and social media information contained conversations about many of these alerts simultaneously.

- Twitter users complained that they received Imminent Threat alerts when their local weather did not exhibit severe weather at all.
- Twitter users expressed confusion at not receiving alerts when others in their network or geographic area did.

3.1.2 AMBER Alerts

The following list contains the results of the analysis of online and social media sources for AMBER alerts using commercial and HSSEDI research tools.

- Twitter users expressed frustration over individuals posting photos of a missing child without including additional information from the corresponding AMBER alerts.
- Commentary on Twitter highlighted the discrepancy between a news source's description of a suspect's vehicle and the description in the AMBER alert.
- Twitter users shamed other users for disabling AMBER alerts on their mobile devices.
- Twitter users expressed frustration over receiving AMBER alerts in locations that were far from the areas mentioned in the alert.
- Although some AMBER alerts were only sent to one state, the conversation online quickly spread beyond state borders, sometimes rising to a national conversation.
- Commentary on Twitter related the success of finding a missing child directly with the EAS.

3.2 Advanced Analytics

The analysis of sentiment in tweets pertaining to the selected WEA messages was based on a set of 396 geo-located tweets (latitude and longitude coordinates from the users' device) between March 2013 and September 2014. The geo-located Twitter data was compared to weather alerts over the same time period.

Tweets were selected around the time of the alert. The distance from the tweet's location to the centroid of the alert's polygon was calculated. It was assumed that the centroid for the weather alert would be approximately where the incident would occur. The tweets were partitioned by distance: within the alert polygon, and 20, 40, and 80 miles from the alert center. This led to 53, 35, 48 and 137 tweets in each distance group, respectively. The remaining tweets were greater than 80 miles away and were not analyzed.

The instances¹⁰ of negative sentiment¹¹ were averaged over tweets in each group and normalized by the number of tweets in each month. The figure below shows a threshold (between 40 and 80 miles) in distance from the alert event where public sentiment becomes much more negative.

¹⁰ See appendices for specific examples.

¹¹ Based on Tweet Explorer's Linguistic Inquiry and Word Count (LIWC) algorithm for detecting negative and positive emotion.

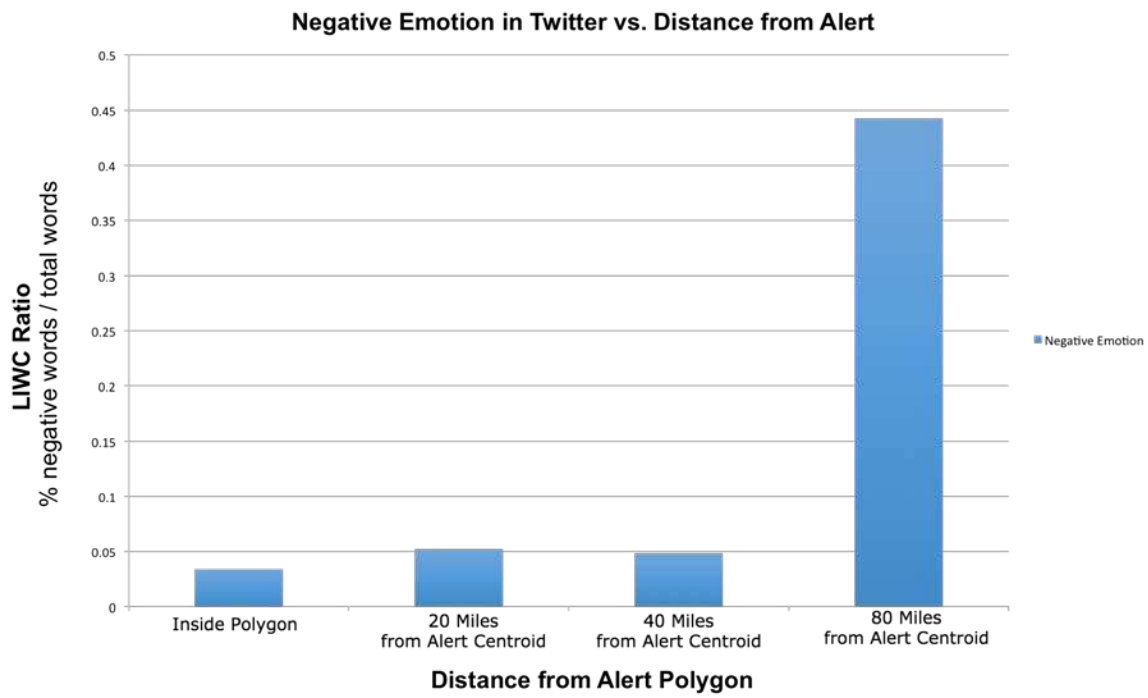


Figure 4: Threshold in Distance for Negative Public Sentiment

4 Observation at NWS

HSSEDI observed NWS's use of social media at a NWS Forecast Office to inform its outreach to the public during Imminent Threat events. During the observation,¹² a tornado warning¹³ was issued and Facebook and Twitter were monitored to gain an understanding of how the public reacted to the alert. At this particular NWS Forecast Office, the staff were already monitoring their Facebook account for user comments.

4.1 Overall Information

Although as many as 70 percent¹⁴ of tornado warnings do not result in a touchdown and when they do occur they usually effect only at the neighborhood level, NWS wants to ensure that the public understands this, takes warnings seriously and responds appropriately. NWS's main mission is forecasting weather and issuing warnings for weather events; the daily workload is often heavy. This means there is limited time to monitor online and social media data sources.

The observed local office is currently using the free version of the commercial tool HootSuite to follow relevant public sentiment on Facebook and Twitter. In this office, Facebook has been monitored for 2-3 years, and Twitter for one year. They use keywords to watch trends related to: general weather, winter weather, severe weather, flood search, etc. This office is also evaluating a Social Media Dashboard, which is being built by another local office.

NWS forecasters have various levels of comfort in using and understanding online and social media sources. Therefore, the use of online and social media data during event monitoring is variable across the 122 NWS Forecast Offices.

4.2 Specific Recommendations for NWS

Based on HSSEDI's observations and analysis, NWS Forecast Offices have various levels of comfort in using and understanding online and social media sources. This is normal in today's work environment; the amount of social media tools is vast and the degree of comfort for each tool varies greatly among employees. This leads to various levels of use of online and social media data during event monitoring across the 122 NWS Forecast Offices. HSSEDI recommends selecting a specific set of tools to rapidly facilitate understanding public sentiment towards specific warnings. This would provide value for local forecasters, while minimizing the additional work load.

HSSEDI recommends leveraging existing trained weather spotters to address additional workload of using social media tools. The observed NWS Forecast Office has 6,000 spotters within its region. These spotters could help monitor online and social media. They could send reports to the respective NWS Forecast Offices directly, or maintain dedicated Twitter or Facebook accounts to communicate with the public. Spotters and other volunteers could provide feedback on whether those subscribers who received the alert were actually in the severe weather

¹² NWS Sterling, Virginia on November 19, 2014.

¹³ This was an actual tornado warning.

¹⁴ Mike Gerber, NOAA Headquarters.

area. This feedback could be crucial in effectively updating an alert or understanding why a particular geographic location was not receiving alerts.

5 Recommendations

HSSEDI's analysis led to two recommendations on how negative public sentiment could be addressed: improved user education through increased information about WEA messages and AO engagement to clarify information related to the events using social media accounts. WEA is still not fully understood by the general public, as evidenced by confusion surrounding why the public gets certain alerts. Education programs need to be expanded to better inform the public about WEA. Additionally, AOs should engage more with the public using social media tools, such as Twitter and Facebook, to clarify any confusion after a WEA message is issued. As part of a general engagement strategy, HSSEDI offers specific recommendations for the NWS based on HSSEDI's observations. Specifically, the NWS should use tools to analyze public sentiment and use existing trained weather spotters to assist with understanding and responding to public sentiment in real time.

Separately, HSSEDI recommends an engagement methodology be used to monitor and respond, via social media, to public sentiment towards WEA. HSSEDI created an initial engagement methodology for AOs to follow using a combination of research and commercial tools. This initial "engagement methodology" could be improved upon and suggestions are provided in Section 4.3.

5.1 General Recommendations for WEA

The following general recommendations address negative public sentiment. These recommendations involve outreach to the general public to better educate them about WEA messages, and more direct engagement with the public online.

5.1.1 Education

Some of the issues noted in HSSEDI's observations can be remedied through education about WEA messages to improve public knowledge about them. Although FEMA has conducted some outreach efforts, such as public service announcements, it is clear through this analysis that there needs to be an increased effort to explain to the public what WEA messages are and why they are important. Some of this effort could be led by the AOs themselves. For example, NWS could sponsor a public education campaign to explain when and why weather alerts are sent. The NCMEC could explain the importance and impact of AMBER alerts. A more informed public could potentially reduce the number of instances when people:

- Express dismay at the loudness or tones made by their cell phones by an alert;
- Complain about receiving weather alerts when there was no evidence of bad weather in their locations;
- Complain about receiving too many alerts of any kind; and
- Wonder why they received no alerts while other mobile users in the same area did.

5.1.2 Engagement

In addition to better educating the public, AOs should more consistently engage with the public to clarify any misperceptions about the WEA messages issued. HSSEDI's analysis shows this engagement to be very sporadic and inconsistent, and is completely dependent on the technical knowledge of the various AOs. Specifically for the NWS and NCMEC, HSSEDI recommends they maintain Twitter accounts, manned 24/7, that would enable near real-time communication using Facebook and Twitter to respond to the public to clarify inaccurate or incomplete information.

Consistent engagement after the WEA message has been issued will help reduce the number of instances, and potentially the severity, of negative feedback.

5.2 Engagement Methodology

HSSEDI developed an engagement methodology for how an AO might use both commercial (labeled industry tools in the figure below) and research tools (labeled custom system in the figure below) to improve public sentiment to WEA messages. This methodology involves a six step process. In the figure below, the steps for the AO are on the left, and the uses for tools are on the right. The methodology involves an AO sending an alert to warn the public of an imminent threat. After alert dissemination, an analyst uses commercial tools to search for and monitor public sentiment. Usually, this will entail follow-up searches using geographical filters. The results of these searches will contain some analytics. The resulting data from the geographic searches is transferred to the custom system for advanced analysis. The results of the advanced analysis contain a synopsis of public sentiment. The analyst will then pass this information along to the AO, potentially enabling future messaging to be more favorably received.

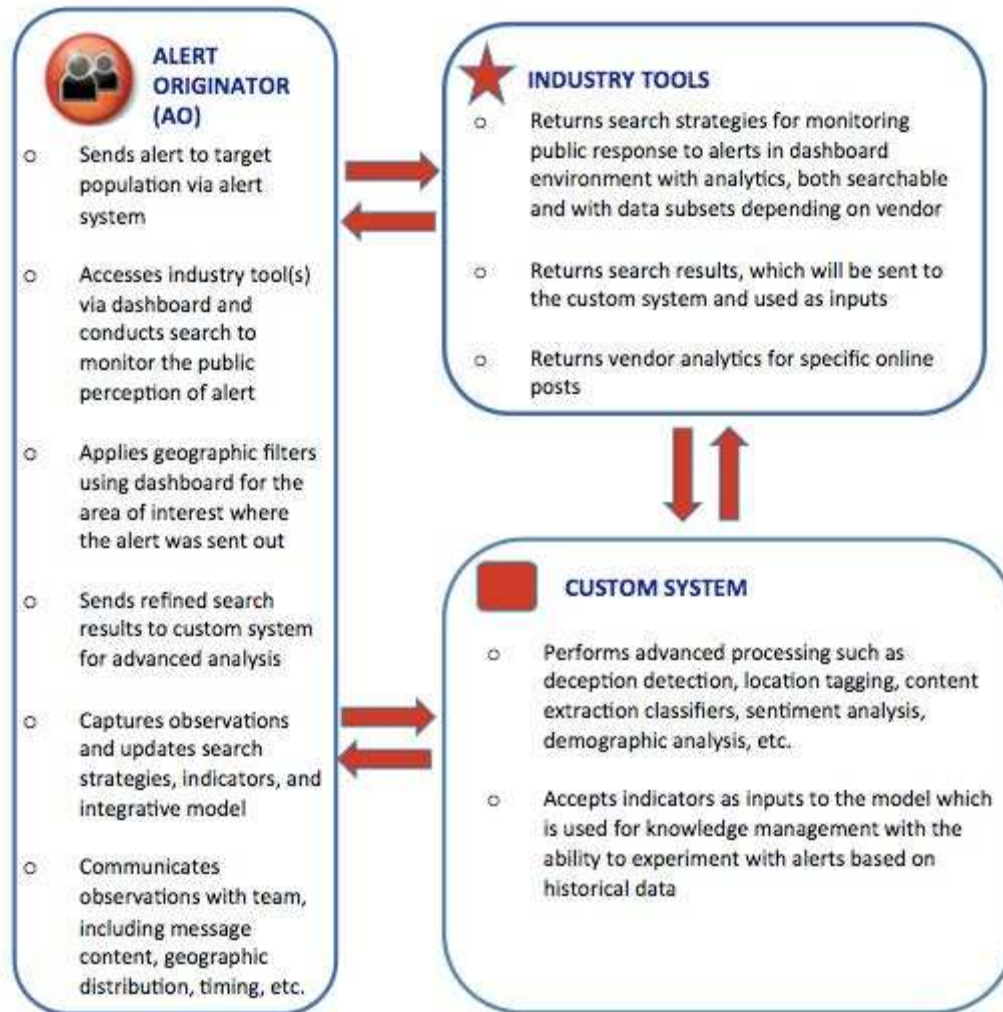


Figure 5: Initial Engagement Methodology

5.2.1 Methodology Improvements

After reviewing the process and results produced, HSSEDI identified improvements that could be made to this initial methodology to better determine sentiment towards WEA messages. The sentiment analysis could take into account new parameters for searches and analysis, such as what county the Twitter user was in, the time of day the alert was sent and the number of alerts received for a given weather event. Additionally, engaging with industry to add much needed capabilities into their tool sets will be extremely important for the long term use of sentiment analysis to gauge the effectiveness and perception of WEA. Training and utilizing multiple analyst roles when using social media to engage with the public would improve efficiency for an AO when either conducting sentiment analysis or responding to public comments. Finally, leveraging volunteers, such as a Community Emergency Response Team¹⁵ (CERT), would help AOs take full advantage of social media analysis.

¹⁵ <https://www.fema.gov/community-emergency-response-teams>

5.2.1.1 Improved Search Parameters

As an example of including new parameters during searches, “AMBER alert” is sometimes used as slang or in a joking manner among online posters. One could develop a classifier (new search parameter) to help isolate these types of posts and remove them from the data set. This will make the resulting data set more accurately reflect the actual public sentiment about WEA. Including other search parameters, such as time of day of the WEA message and corresponding public comment, could also improve search results. Limiting the timespan after which a WEA message was issued potentially could provide meaningful insights into public sentiment. Comments made within six hours of a WEA message being issued may have more weight or meaning than comments made six days after the message was issued. AOs should spend some effort in developing their search criteria to make their search results more meaningful.

5.2.1.2 Industry Engagement

Sentiment analysis, improved geo-location, clustering and meme detection are other advanced analytic capabilities that may provide value. As these capabilities are currently not available in commercial tools, engaging with industry to include these capabilities would be beneficial for AOs. While these capabilities could be developed in custom-built research tools, AOs should limit reliance on such tools. Often research tools are not supported to the same degree as commercial tools in both new capability development and user support. Commercial vendors have a vested interest in ensuring that their tools work correctly and meet their customer’s needs. When possible, AOs should engage with industry to improve the commercial offerings.

5.2.1.3 Increase Efficiency

Some strategies for increasing efficiency include adequate training and multiple analyst roles. Using tools to monitor public sentiment will require training and knowledge of best practices. Relying on employee familiarity with social media and social media tools will yield inconsistent results at best.

Developing a training program to educate employees on proper usage of social media and social media tools is essential to ensuring consistent results. Additionally, splitting the analyst role into two roles, collection analyst and operations analyst, may help improve efficiency, as each analyst can focus on the tasks for the defined role. Collection of data, even using a custom system, requires an analyst to possess detailed knowledge of a search strategy, contextual knowledge of the type of data that strategy is aimed at finding, and knowledge of filtering techniques to obtain the best dataset. An operations analyst requires a slightly different methodology, including experience with advanced analytics and how each analysis capability contributes to measuring public sentiment.

5.2.1.4 Leveraging Volunteers

As discussed in the specific recommendations for the NWS, leveraging volunteers, such as a CERT or weather spotters, could potentially help AOs take full advantage of social media analysis. These volunteers could feed meaningful information back the AO, helping free the AO from gathering public comment. Ultimately, the AO’s job is not to analyze public sentiment about their WEA messages, but to ensure that the proper information gets to the public. Volunteers could potentially fill the niche and support the AO by conducting sentiment analysis for them, or at least pre-filtering information.

6 Conclusion

Social media sentiment analysis is an emerging capability. No single commercial tool offers all of the needed capabilities to provide the kind of meaningful information needed by AOs. Custom-built research tools fill in some of the missing capabilities, but even these are not full featured. Knowledgeable analysts are required to use a variety of tools to feed output from one tool to the input of another so meaningful results can be gathered. The recommendations in this report provide a way forward for AOs to start to gather public sentiment about WEA messages and help improve the sentiment about these messages. As the public becomes more knowledgeable and less suspicious of WEA messages, public sentiment towards them will improve. AOs should be cognizant of this when determining how much time and effort they want to put into examining public sentiment and responding to public comment via social media.

Ultimately, sentiment analysis will only be as useful and accurate as the available analysis tools and analyst training allow. The long term success and usage of sentiment analysis will be dependent on maintaining trained staff and engaging with industry. Although natural language parsers and sentiment analytics are improving daily, the complexities and nuances of speech and slang used, particularly in social media,¹⁶ will necessitate using analysts to fill the gaps where sentiment software falls flat. Trained analysts will lead to more consistent results and meaningful data.

¹⁶ <http://blog.hubspot.com/marketing/how-internet-changes-language>

Appendix A Historical Analysis

The social media sources related to Wireless Emergency Alerts (WEA) were analyzed using commercial and Homeland Security Systems Engineering and Development Institute (HSSEDI) research analysis tools, for both weather and America's Missing: Broadcast Emergency Response (AMBER) alerts. Lessons learned are defined as patterns observed by reading the content of the public's response to WEA message types. The detailed analyses below illustrate how the Alert Originator (AO) could benefit from this type of analysis in understanding the effectiveness of alerts sent.

A.1 Analysis Methodology

To generate useful analysis, relevant data must be obtained. "Relevant" means data pertaining to public sentiment of imminent threat weather or AMBER alerts, while minimizing tweets on other topics from appearing in the data set. Determining what to search for online and within social media sources is necessary before analysis can be conducted; this is referred to as a search strategy.

Search strategies are composed of keywords from background material on the topic of interest and used with Boolean search to obtain data from online and social media sources. A subject needs to be at least cursorily investigated at the beginning of the effort to choose the right entity names to include in a strategy including: user names, organization names, geographical names, topic-specific terminology, appropriate expressions, hashtags (subjects in Twitter denoted with #), slang and/or jargon. This allows a data set to be collected that is associated with the WEA messages during specific date ranges. Keywords are combined using Boolean operators or proximity operators to elicit incidences of the words on the topic of interest and to exclude ambiguous terms or "noise" (e.g., looking for: ("dengue" or "fever") AND NOT ("Bieber" OR "spring" OR "boogie")).

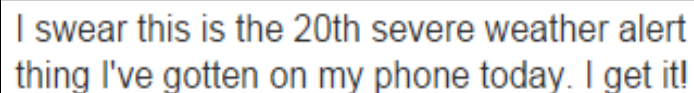
Creating a search strategy relies on iterative testing, adding and removing keywords to see what sort of tweets are returned. The analyst can do this using the public Twitter website or an industry. One iteration of a search strategy is never enough; as search results inform the construction of subsequent strategies and events on the ground change over time, new keywords emerge and are dropped for the topic at hand. With specific search strategies developed for imminent threat weather and AMBER alerts,¹⁷ the analyst is ready to obtain the data.

¹⁷ **Weather Alert Query:** (("weather alert" AND "text") OR ("weather alert" AND "message") OR ("weather alert" AND "phone") OR ("weather alert" AND "mobile") OR ("tornado alert" AND "text") OR ("tornado alert" AND "message") OR ("tornado alert" AND "phone") OR ("tornado alert" AND "mobile") OR ("flood alert" AND "text") OR ("flood alert" AND "message") OR ("flood alert" AND "phone") OR ("flood alert" AND "mobile") OR ("avalanche alert" AND "text") OR ("avalanche alert" AND "message") OR ("avalanche alert" AND "phone") OR ("avalanche alert" AND "mobile") OR ("storm warning" AND mobile) OR ("storm warning" AND phone) OR ("storm warning" AND message) OR ("storm warning" AND text) OR ("dust storm" AND mobile) OR ("dust storm" AND message) OR ("dust storm" AND phone) OR ("dust storm" AND text) OR ("extreme wind" AND mobile) OR ("extreme wind" AND phone) OR ("extreme wind" AND text) OR ("extreme wind" AND message))

AMBER Alert Query: (("amber alert " AND "text") OR ("amber alert" AND "message") OR ("Amber alert" AND text) OR ("Amber alert" AND message) OR ("amber alert" AND "phone") OR ("amber alert" AND noise) OR ("Amber alert" AND loud) OR ("amber alert" AND "mobile"))

A.2 Tornadoes and Flooding in Arkansas and Mississippi

From April 27-30, 2014, tornadoes and flooding swept through Arkansas and Mississippi, prompting 20 tornado and flash flood alerts in the two states. The alerts garnered attention and chatter in online and social media. Posters commented on the repetition and volume of the alerts (Figure 6), while one news article attributed the low loss of life (one death) to the alert system: “I think the reason that there was only one loss of life in a community like this was because of the emergency response or the emergency warning system.”¹⁸ Top memes (defined as ideas in text) found in the data reflect message content of storm warnings, instructions on how to enable a Red Cross alert on a mobile phone, and retweets about tornado, flood and dust storm warnings for specific counties as shown in Figure 7. Note the meme is on the left, followed by the number of tweets composing the meme, the data range of the tweets and a link to go to the individual tweets. A near real-time meme detection system can help users track topics and read the online posts in an efficient manner. Figure 8 shows the geographic distribution of alerts for this study. Being able to geospatially broadcast alerts and understand where online posters are when they have an issue is a key component to understanding what is happening on the ground. Locating more online posts is a necessary first step. Current research algorithms can locate online posts down to the county level.



I swear this is the 20th severe weather alert thing I've gotten on my phone today. I get it!

Figure 6: A Tweet Expresses Annoyance at Large Number of Alerts

¹⁸ Trujillo, M. (November 18, 2013) Schock: Warning System Saved Lives. <http://thehill.com/video/house/190551-schock-warning-system-saved-lives>.

| MemeME | | | | |
|---|----|---|---------------------------|------------------------------|
| RT @DiannaGee: HOW TO: Set An Alert for Tornado Warnings On Your Phone http://t.co/rsy2dHt7y3 | 13 | 2014/04/27 14:20:12 -0400 | 2014/04/28 17:48:05 -0400 | Go to Tweets |
| iNWS Hydro Alert - AHPS Mobile Alert Dan River at South Boston (SBNV2) A CHANGE IN FLOOD CATEGORY HAS BEEN... http://t.co/KsHbmQ2C3Q | 11 | 2014/04/28 22:40:18 -0400 | 2014/04/29 22:20:38 -0400 | Go to Tweets |
| iNWS Hydro Alert - AHPS Mobile Alert Dan River at Paces (PCEV2) A CHANGE IN FLOOD CATEGORY HAS BEEN... http://t.co/sUaSJmJhww | 5 | 2014/04/29 08:20:47 -0400 | 2014/04/29 22:20:34 -0400 | Go to Tweets |
| Flash Flood Warning for Mobile and Baldwin Counties, Alabama - http://t.co/XOe2LsY429 Latest Flood Warning Update! | 4 | 2014/04/29 16:45:00 -0400 | 2014/04/29 22:37:01 -0400 | Go to Tweets |
| RT @TROYUSOS: Please stay alert to changing weather. Check email, text and http://t.co/AiCihqdQ4v for updates on weather and University op... | 4 | 2014/04/28 17:39:35 -0400 | 2014/04/28 19:44:35 -0400 | Go to Tweets |
| RT @SpencerWeather: Safety being taken seriously. GOOD! RT: "@mfoxsellers: #noaa tornado alert going off on my phone. @wsfa12weather http://t.co/... | 3 | 2014/04/29 03:23:33 -0400 | 2014/04/29 04:43:45 -0400 | Go to Tweets |
| RT @WheresAndrew: Just received this alert in my phone. Wind gusts @ 55mph on the road & dust storm warnings. You're not easy, Texas. http://t.co/... | 3 | 2014/04/29 11:58:31 -0400 | 2014/04/29 14:40:59 -0400 | Go to Tweets |
| RT @FunnyMaine: if the tornado doesnt give you a heart attack, the loud ass phone alert will | 3 | 2014/04/29 17:59:22 -0400 | 2014/04/29 19:13:04 -0400 | Go to Tweets |
| Showing 1 to 10 of 16 entries | | | | |
| | | Previous Next | | |

Figure 7: Grouped Ideas or Memes



Figure 8: Geographic Distribution of Tweet Postings

A.3 Severe Weather in Texas and Other States

From October 4-21, 2013, a number of weather alerts affected Texas and other states. Analysis showed two distinct events for this data set, one during early-October and the other during mid-October as shown in Figure 9. These peaks were individually sent to Comment Filter (CoFi) for machine automated clustering. Although analysis of tweets during both peaks reveal similar conversations about “flash floods,” the data sets differ in the states and counties mentioned. During the first peak, the conversation is dominated by flash flood warnings in Kentucky, while the second peak’s conversation was dominated by flash flood warnings in Texas and Colorado as shown in Figure 10 and Figure 11, respectively. This tool could automatically cluster data in near-real time (left column automatically labeled “my machine”) so users could understand the conversation and public sentiment around WEA messaging.

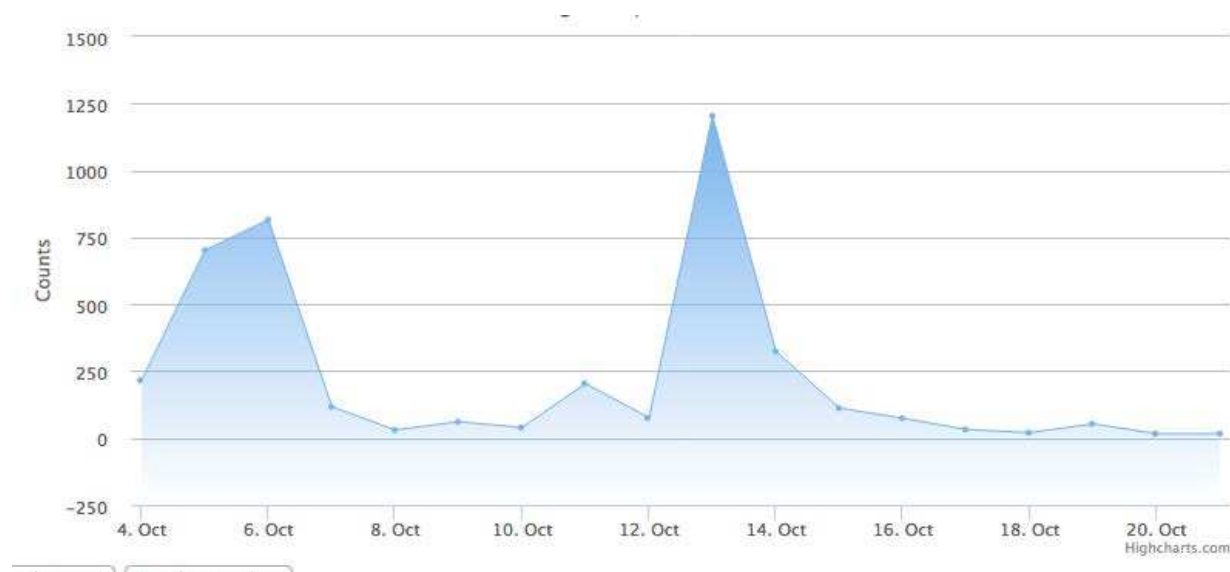
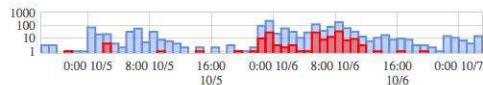


Figure 9: Two Peaks in Early and Mid-October



Filter Text:
Filter Users:

Cluster Keywords: **ky until, in ky, edt sunday, jefferson county, for jefferson**; Cluster size: **165 comments (11.8%)**

☐ Select all ☐ Export Selected

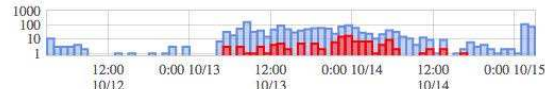
Cluster Key

x Overview

county, due, warning issued
ky until, in ky, edt sunday
county, service in
issued for, effect for
my phone, warning on
congress, service employees
october 05, cdt until
by nws, edt until, edt by
cdt sunday, tristatewx
nws has, has issued
scared the, my phone
pm http, under flash, 00 pm
fillmore houston, 00am
saturday
of rain, inches of
from 10, 10 2013
floyd howard, howard
mitchell
cdt expiring, edt expiring
eastern, fox6weather
next hours, tonight http
45 many, warning extended
View All Comments

☐ 1266, 2013-10-06T07:58:24Z / **SkyTrackWX**: Flash Flood Warning for Jefferson County in KY until 8:45am EDT Sunday. #KYwx
☐ 785, 2013-10-06T02:13:38Z / **BullittPHEM**: RT @KevinHarned: Flash Flood Warning for Jefferson County in KY until 2:00am EDT Sunday. http://t.co/U3EYzvXZ8V
Previous comment repeated 3 additional times.
☐ 1173, 2013-10-06T05:59:32Z / **SkyTrackWX**: Flash Flood Warning for Jefferson County in KY until 7:00am EDT Sunday.
Previous comment repeated 1 additional times.
☐ 768, 2013-10-06T02:10:59Z / **KevinHarned**: Flash Flood Warning for Jefferson County in KY until 2:00am EDT Sunday. http://t.co/U3EYzvXZ8V
☐ 767, 2013-10-06T02:10:58Z / **WAVE3Weather**: Flash Flood Warning for Jefferson County in KY until 2:00am EDT Sunday. #wave3weather
☐ 1168, 2013-10-06T05:58:43Z / **wave3news**: RT @KevinHarned: Flash Flood Warning for Jefferson County in KY until 7:00am EDT Sunday. http://t.co/U3EYzvXZ8V
☐ 1296, 2013-10-06T10:31:10Z / **BenJordan88**: RT @KevinHarned: Flash Flood Warning for Jefferson County in KY until 8:45am EDT Sunday. http://t.co/U3EYzvXZ8V
Previous comment repeated 1 additional times.
☐ 1262, 2013-10-06T07:56:25Z / **KevinHarned**: Flash Flood Warning for Jefferson County in KY until 8:45am EDT Sunday. http://t.co/U3EYzvXZ8V
☐ 1157, 2013-10-06T05:57:36Z / **KevinHarned**: Flash Flood Warning for Jefferson County in KY until 7:00am EDT Sunday. http://t.co/U3EYzvXZ8V
☐ 1156, 2013-10-06T05:57:35Z / **WAVE3Weather**: Flash Flood Warning for Jefferson County in KY until 7:00am EDT Sunday. #wave3weather
☐ 1261, 2013-10-06T07:56:24Z / **WAVE3Weather**: Flash Flood Warning for Jefferson County in KY until 8:45am EDT Sunday. #wave3weather
☐ 772, 2013-10-06T02:11:18Z / **JohnBersotjr**: RT @WAVE3Weather: Flash Flood Warning for Jefferson County in KY until 2:00am EDT Sunday. #wave3weather
☐ 1108, 2013-10-06T06:01:31Z / **KentuckyWX**: RT @WAVE3Weather: Flash Flood Warning for Jefferson County in

Figure 10: CoFi Analysis of First Peak Yields Kentucky Conversation



Filter Text:
Filter Users:

Cluster Keywords: **in tx, warning for, for dimmit, county in, zavalas**; Cluster size: **116 comments (7.4%)**

☐ Select all ☐ Export Selected

Cluster Key

x Overview

warning for, issued
cdt until, issued october
the flash, issued, day, acl
my phone, scared the
for travis, williamson
counties
in tx, warning for
details http, rt
txstormchasers
national weather, the
national
october 14, hst until
simpleweather4u flash,
thunderstorms
sunday http, until 00am
colorado wharton, counties
until
15pm mdt, nm, for lea
oahu until, warning posted
cdt expiring, alert flash
at 00am, 00am cdt, nws the
00am cdt, at 00am, nws http
maverick county, for

☐ 1117, 2013-10-13T21:26:34Z / **KABBFOX29**: Flash Flood Warning for Dimmit and Zavala County in TX until 6:30pm Sunday. http://t.co/14iXXzPIDJ
☐ 892, 2013-10-13T18:36:52Z / **KABBFOX29**: Flash Flood Warning for Dimmit and Zavala County in TX until 4:30pm Sunday. http://t.co/14iXXzPIDJ
☐ 733, 2013-10-13T15:26:05Z / **KABBFOX29**: Flash Flood Warning for Dimmit and Zavala County in TX until 1:30pm Sunday. http://t.co/14iXXzPIDJ
☐ 1118, 2013-10-13T21:26:35Z / **News4SA**: Flash Flood Warning for Dimmit and Zavala County in TX until 6:30pm Sunday. http://t.co/xbpf7biPNO
☐ 893, 2013-10-13T18:36:53Z / **News4SA**: Flash Flood Warning for Dimmit and Zavala County in TX until 4:30pm Sunday. http://t.co/xbpf7biPNO
☐ 734, 2013-10-13T15:26:06Z / **News4SA**: Flash Flood Warning for Dimmit and Zavala County in TX until 1:30pm Sunday. http://t.co/xbpf7biPNO
☐ 894, 2013-10-13T18:37:02Z / **9NYNE9**: RT @KABBFOX29: Flash Flood Warning for Dimmit and Zavala County in TX until 4:30pm Sunday. http://t.co/14iXXzPIDJ
☐ 1178, 2013-10-13T23:32:47Z / **KABBFOX29**: Flash Flood Warning for Dimmit County in TX until 9:30pm Sunday. http://t.co/14iXXzPIDJ
☐ 1179, 2013-10-13T23:32:48Z / **News4SA**: Flash Flood Warning for Dimmit County in TX until 9:30pm Sunday. http://t.co/xbpf7biPNO
☐ 1188, 2013-10-13T23:41:31Z / **rosemary_gomez**: RT @KABBFOX29: Flash Flood Warning for Dimmit County in TX until 9:30pm Sunday. http://t.co/14iXXzPIDJ
☐ 1124, 2013-10-13T21:29:40Z / **BreakinNews01**: RT @News4SA: Flash Flood Warning for Dimmit and Zavala County in TX until 6:30pm Sunday. http://t.co/xbpf7biPNO
☐ 944, 2013-10-13T19:44:27Z / **KABBFOX29**: Flash Flood Warning for Maverick County in TX until 5:30pm Sunday. http://t.co/14iXXzPIDJ
☐ 974, 2013-10-13T19:57:53Z / **KABBFOX29**: Flash Flood Warning for Dimmit, Maverick, and Zavala County in TX until 6:00pm Sunday. http://t.co/14iXXzPIDJ

Figure 11: CoFi Analysis of Second Peak Yields Texas and Colorado Conversations

A.4 Flash Floods in Hawaii

On December 30-31, 2013, multiple flood alerts were sent for some Hawaiian Islands with several peaks in Twitter activity as shown in Figure 12. Breakpoints analysis is described in Appendix C. This graph shows the number of tweets (y-axis) over time. Using CoFi, as shown in Figure 13, the data reflects multiple retweets of weather service flood warnings. The data set was “clean” in the sense that any retweeted weather messages were on topic (not noise); they included either rewordings of the alerts or encouragements to pay attention to the weather warnings for particular areas and evacuate as needed. There was no commentary on loudness of alerts, users annoyed or confused by alerts, other than one individual remarking that the sky was clear blue and there was no sign of downpours or flooding. This is an illustration of a clean data set that shows very positive response from the public.

A drill down into a CoFi message cluster provided these examples:

- 87, 2013-12-30T13:46:47Z / #hiwx Flash Flood Warning for the Hamakua section of the Big Island til 6:30 am. Radar at 3:38 am: <http://t.co/roVOXBxepC>
- 520, 2013-12-31T14:42:42Z : Flash Flood Warning extended until 8am this morning.. just got the Emergency Alert on my phone...
- 259, 2013-12-30T19:08:05Z / : Flash Flood Warning once again up for the Big Island until Noon today. Covers mainly Hilo and Puna areas

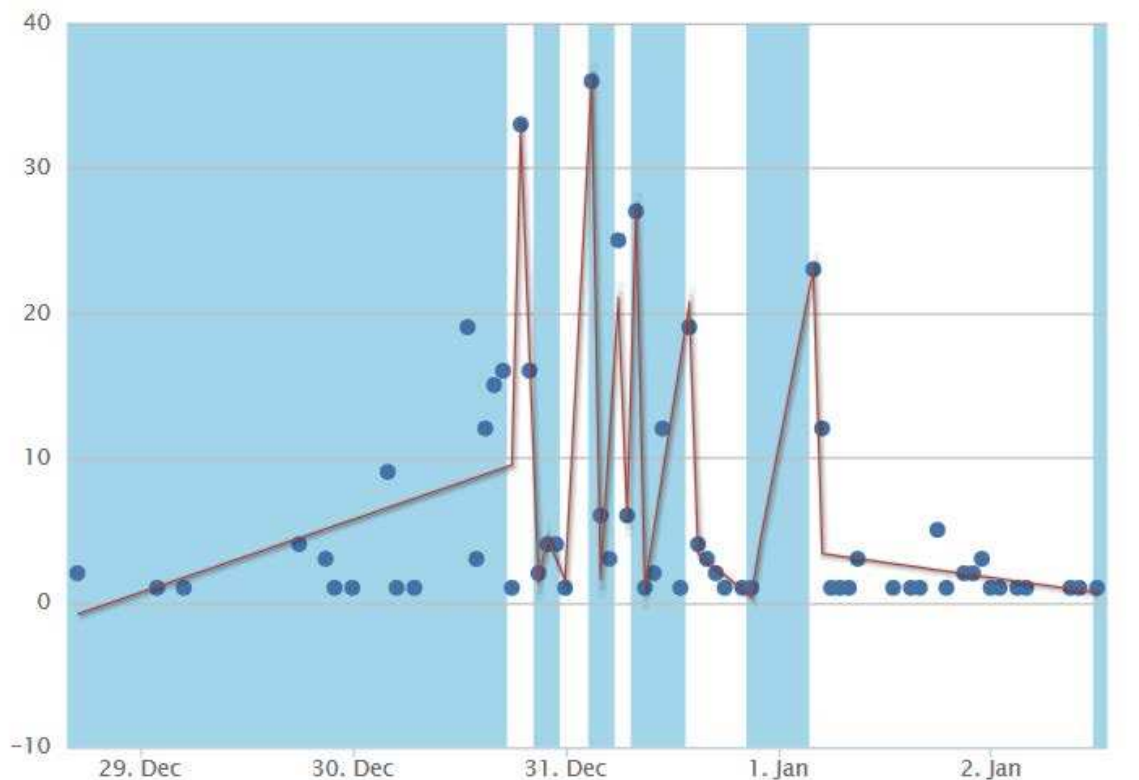


Figure 12: Multiple Breakpoints Over Two Days

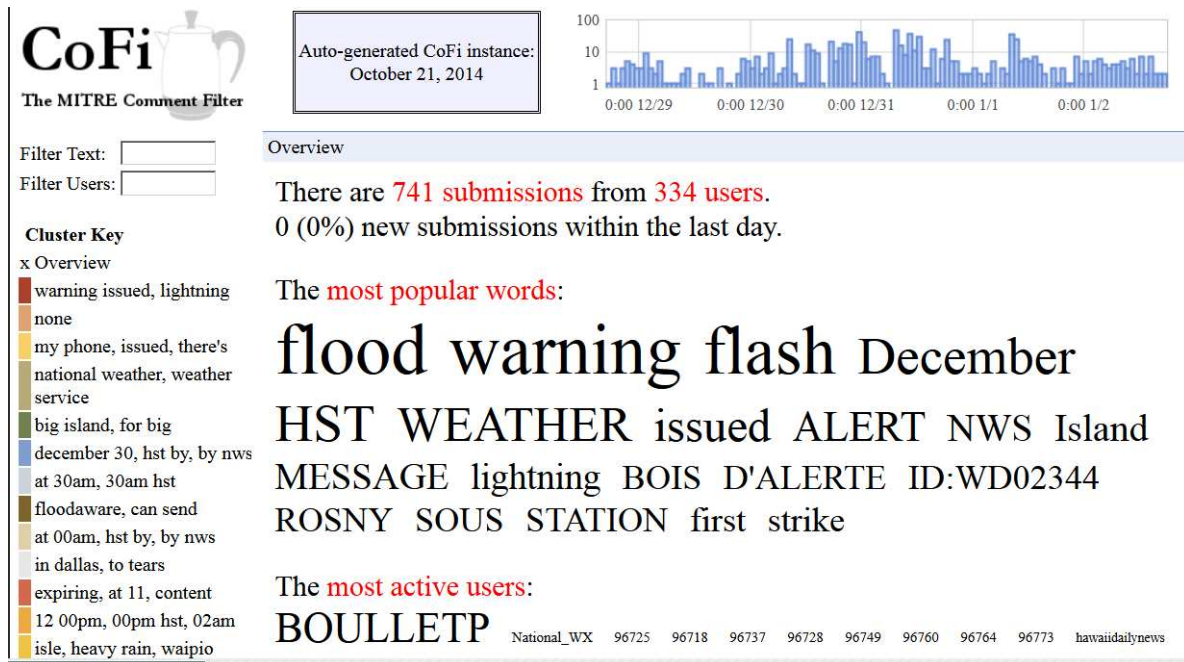


Figure 13: CoFi Analysis of Hawaiian Flash Flood

A.5 Multiple AMBER Alerts in January 2014

During January 2014, six AMBER alerts were sent. Given the overlap in time, there is no way to ungroup them in online and social media posts. This represents a unique challenge and is the reason it was selected as a study. Twitter data was gathered that covers all six cases; a peak occurred on January 11, 2014 as shown in Figure 14. Analyzing tweets in CoFi for January 11, 2014 (Figure 15), showed that Alize Whipple, one of the missing children, was the reason. The most popular words include “Alize,” “Whipple,” “Fitchburg” and “MassStatePolice,” all related to Alize Whipple's case. Popular words also picked up on Alize's mother.

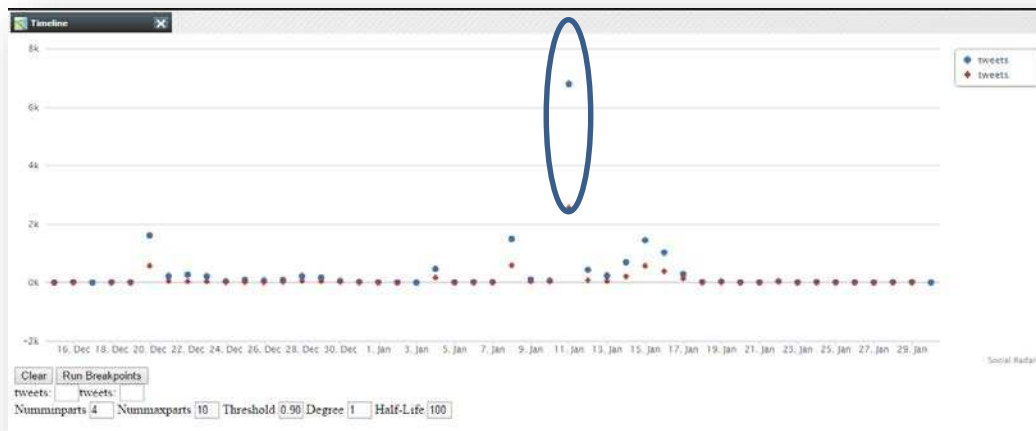


Figure 14: Six AMBER Alerts with a Peak on January 11, 2014



Figure 15: CoFi Analysis Shows Alize Whipple's AMBER Alerts Dominates Conversation

One of the CoFi clusters included a group of Twitter users that were scolding other users who complained about receiving an AMBER alert. Examples include:

- 5910, 2014-01-11T16:18:26Z / *You people are pissed off over an amber alert on your phone?*
- 4017, 2014-01-11T03:41:31Z / *you just might be the only person to complain about getting an amber alert sent to your phone smh*

Another CoFi cluster included Twitter users asking other users if they received an alert in the same area. Examples include:

- 6116, 2014-01-11T18:51:22Z / *did you not get the amber alert on your phone?*
- 3170, 2014-01-11T03:21:38Z / *Hi Jen...since we're in the same area, are you getting that Amber alert on your phone, too?*

Both of these cluster examples represent a point of engagement for AOs. Reaching out to the public with education or engagement from an official Twitter account could help resolve confusion, increase reach or save lives.

Analyzing the larger data set for the entire date range and searching for the names of the AMBER alert children reveals interesting trends. For each child, conversation quickly increased and subsequently decreased as the case came to a close, as shown in Figure 16. Some children generated more conversation than others. According to analysis, this was often a factor of the geographic coverage of the case and the way the case unfolded. Analysis also showed that while

activity peaked and died down for individual AMBER alerts relatively quickly, automatic clustering detected in the data set remained at high levels throughout. These clusters automatically tease out common threads in the data set, such as retweets, the mention of “my phone,” “alerts,” “noise” and even the names of children for those cases that generated more conversation than others. The basis of a near real-time system would need the ability to isolate discussion about individual message sets. This example shows that even when six AMBER alerts were sent in one month, text analytics tools can still isolate them. This will enable AOs to act on specific message sets.

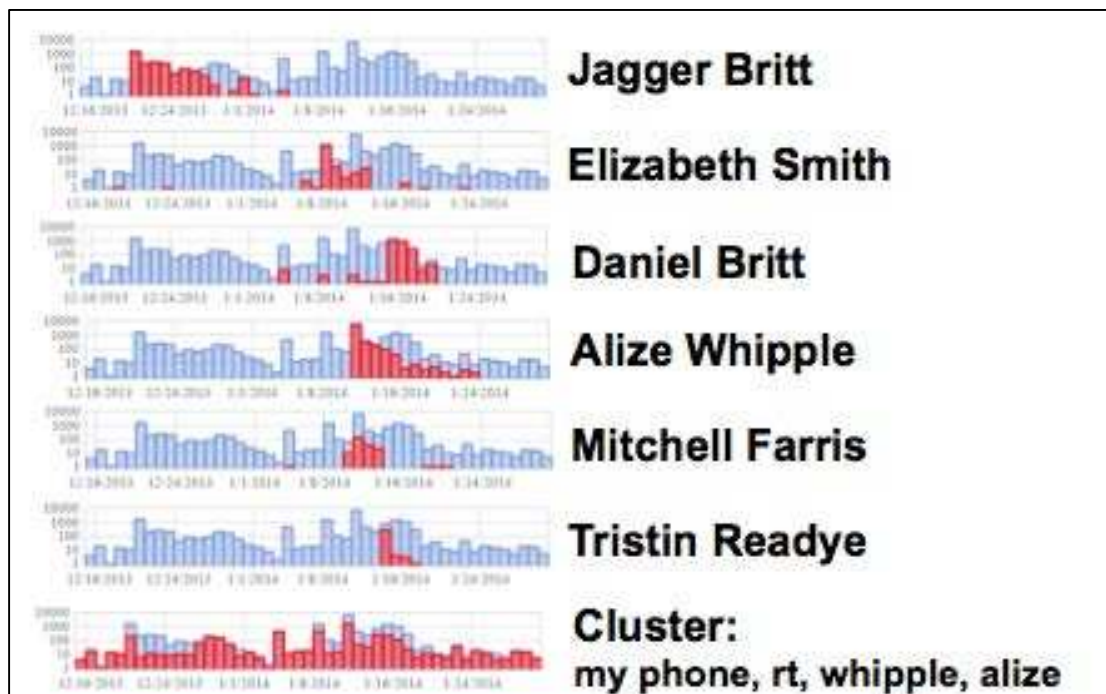


Figure 16: Chatter Increases Quickly then Decreases as Case Status Changed

HSEDI analysts used Pinocchio, a similarity analysis tool, to reveal a network of retweets about another AMBER alert case, Elizabeth Catherine Smith. The tool detected a network of retweets about Elizabeth Catherine Smith. This approach allows the user to drill into the data by network rather than simply cluster. Meme and clusters provide value in isolating the topic of discussion; coordinated account analysis is another method of determining the conversation.

After isolating the topic of discussion, sentiment analysis tools are used to gauge public sentiment. Sentiment analysis for the January AMBER alerts shows a large spike in mid-January in Figure 17. Topics detected during this time include the hashtag #amber, as well as news sources and hashtags of locations where the children were abducted.

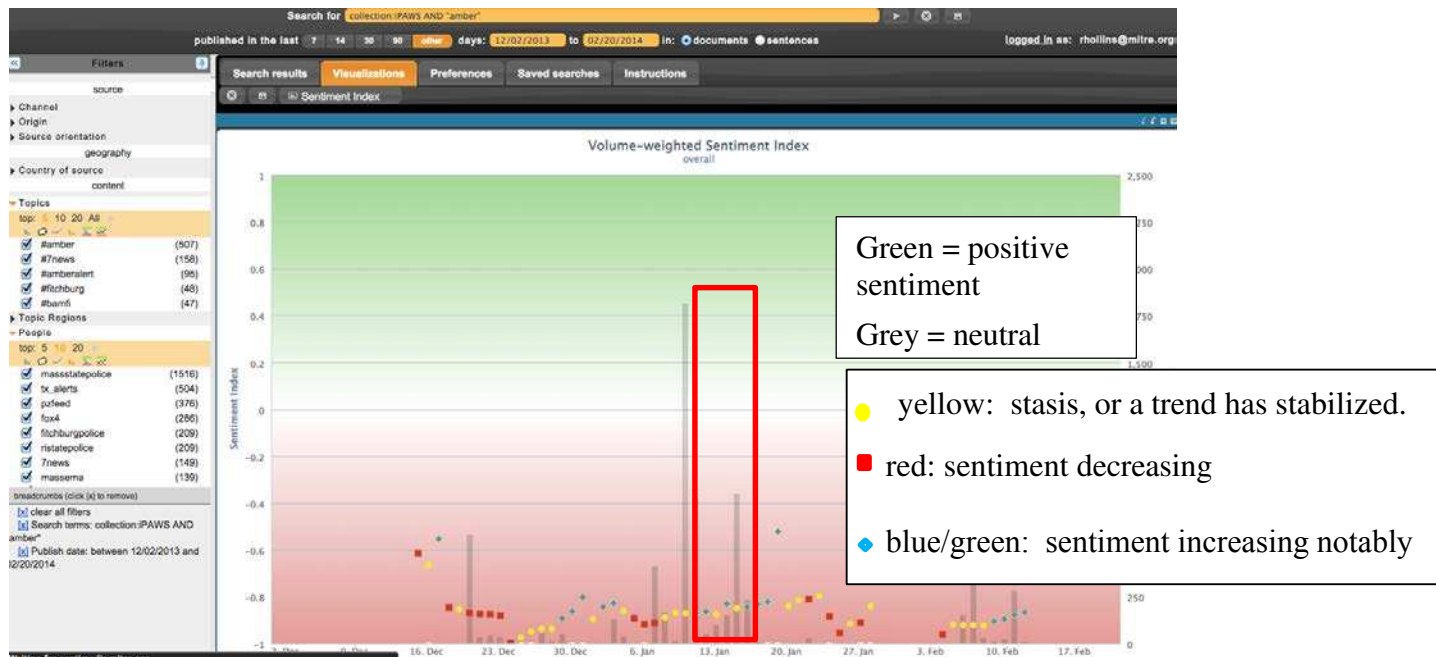


Figure 17: Spike in positive sentiment expressed in messages

The majority of negative tweets during this time mention the frequency of weather alerts, how an alert scared the Twitter user, and one particular user disabling the AMBER alert notifications, as shown in Figure 18.

| #pos | #neg | title | published |
|------|------|--|------------|
| -- | -- | 13 4th flash flood alert on my phone today ☹ | 04/30/2014 |
| -- | 1 | 13 RT @nate_kidd_13: That flood alert thing on my phone just scared the complete out of me. | 04/30/2014 |
| -- | 1 | 13 What the heck? My phone just buzzed that stupid tone emergency alert for a tornado warning. Yeah. Ok. | 04/30/2014 |
| -- | -- | 13 Obviously I turned off the amber alert notifications on my phone so I knew something bad was happening when my phone flipped out #weather | 04/30/2014 |
| -- | -- | 13 I swear this is the 20th severe weather alert thing I've gotten on my phone today. I get it! | 04/30/2014 |
| -- | 1 | 13 RT @darby_michelle: This weather alert on my phone keeps scaring the mess out of me ☹ | 04/30/2014 |
| -- | -- | 13 Flash Flood Warning for Mobile and Baldwin Counties, Alabama - http://t.co/MtaakNQJ5w phone locked up warning came out about 45 min ago. | 04/30/2014 |
| -- | 1 | 13 Just about had a heart attack from that flash flood alert on my phone | 04/30/2014 |
| -- | 1 | 13 That flood alert thing on my phone just scared the complete out of me. | 04/30/2014 |
| -- | -- | 13 Few things are spookier than being in a quiet store and everyone's cell phone weather alert going off simultaneously. | 04/30/2014 |
| -- | -- | 13 RT @niniehammon: Accidentally left my cell phone on during a flight and got a flash flood alert. Figured I was safe. | 04/30/2014 |
| -- | 1 | 13 That weather alert just scared me. I forgot how loud that alarm sounded on my phone. | 04/30/2014 |

Figure 18: Drilldown Using Sentimeditr

The final analysis output, shown in Figure 19, is the relative proportions of sentiment categories using Crimson Hexagon. Negative categories dominated the conversation with mentions of alert loudness followed by complaints about frequency of alerts or the time the alert was sent.

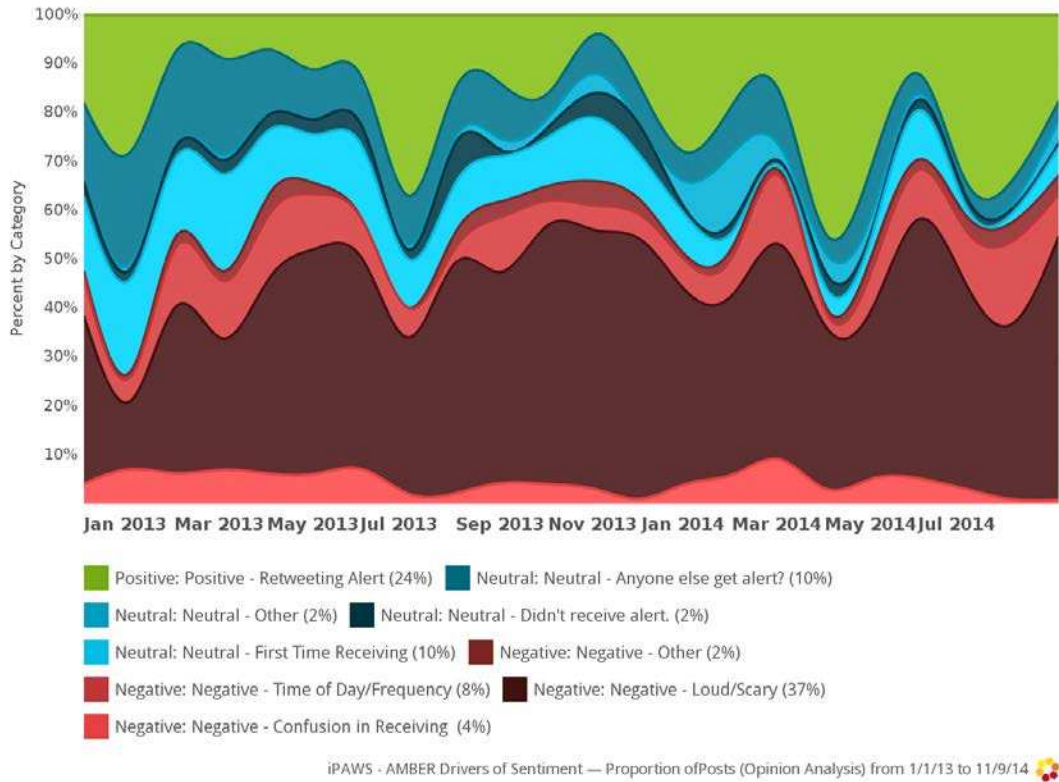


Figure 19: Sentiment Analysis from Crimson Hexagon

Appendix B Commercial Tools

The following is a list of commercial tools. HSSEDI does not endorse any tool or vendor. HSSEDI's assessment is ongoing and is not exhaustive. In comparing each type of commercial tool, Twitter analytics is used as a benchmark. Commercial tools are broken down into five tool categories listed in order from most sophisticated to least sophisticated:

1. Full Twitter Search Dashboard Capability with One Year Archive
 - a. Marketwired Sysomos MAP
2. Full Twitter Search Dashboard Capability with One Month Archive
 - a. InTTensity
 - b. Salesforce Radian6
3. Full Twitter Search Capability and Results Brought into Dashboard
 - a. Crimson Hexagon ForSight (four-year data archive)
 - b. Babel Street
 - c. Lockheed Martin (LM) WISDOM / DataSift
 - d. Nuvi
 - e. Marketwired Sysomos Heartbeat
 - f. Recorded Future including Forecasts
 - g. SAS Social Media Analytics
 - h. Spiral16

- i. Synoptos Public
 - j. Twitter Gnip App
- 4. Full Twitter Search Capability and Results Brought into Tool with Specialized Analytics
 - a. Geofeedia
 - b. Graphika
 - c. HumanGeo's Media Monitoring
- 5. Full Twitter Search Capability with Dashboard Streaming
 - a. Hashtagify.me
 - b. HootSuite
 - c. Meltwater Buzz
 - d. NetVibes
 - e. Spredfast
 - f. SocialMention
 - g. SocialOomph
 - h. SproutSocial
 - i. TweetDeck

Commercial tools vary widely in the features they provide and the usefulness of those features. For example, InTTensity and Radian6 are multisource analytic dashboards that provide only one month of historical data. In contrast, a few products allow search back for multiple years instead of only one month (e.g., Crimson Hexagon ForSight, Marketwired Sysomos MAP).

The tools listed under “Full Twitter Search Capability with Dashboard Streaming” are not as sophisticated as some of the other analytic tools, but can be used to schedule tweets or other social messages in bulk (same messages going to multiple accounts at pre-determined times) and subsequently gauge reaction to these messages. This is particularly useful for public messaging campaigns. Alerts can be defined for key terms that will then notify the user when those occurrences appear.

SocialMention offers a sentiment detection (positive, neutral, negative) capability, and most of the tools have “engagement” options for reaching out to the public. Engagement can be driven by volume of response and reaction to a particular theme or post, which may result in reposting of the message or rebuttal to content.

Hashtagify.me is somewhat different than other tools in the list; it is a near real-time tool that is useful for the purpose of identifying keywords and hashtags that can be leveraged in other tools. As hashtags change frequently, messaging and alerts can be modified to contain or monitor for them.

B.1 Babel Street

Founded in 2009, Babel Street is a privately owned social media monitoring firm based in Reston, Virginia. Users create their own searches in either the basic or advanced search interface. The tool is typically deployed as a web application on the Amazon Web Services cloud, although it may be deployed on a customer site as a separate instance. Babel Street reports that most users typically search only the present day and do not use historical data; however, 1-4 years of the full Twitter fire hose is available. Sources are typically segmented by user accounts and users are matched with a source profile that best suits their needs. Babel Street is not a reseller of Twitter data, but exports of tweets can be managed through a license purchase from Twitter and the data

pipled through the Babel Street interface. Babel Street can provide a 1 percent random Twitter feed and can index the data for users. Babel Street has application program interfaces (API) that can handle large amounts of data for pulling streams.

| | |
|--------------|---|
| Babel Street | Licensing is offered at three tiers: the Bering (\$1,250 per month for 100,000 tweets per month (English only), no Application Programming Interface (API) included; the Atlantic (\$2,000 per month for 350,000 tweets per month (five languages), includes API; and the Pacific (\$3,050 per month for 1 million tweets per month (200+ languages), includes API. |
|--------------|---|

B.2 Crimson Hexagon

- Crimson Hexagon’s ForSight platform is a social media analytics tool that mines blogs, Facebook, Instagram, Twitter (full fire hose back to July 2010) and other social media sites to identify social engagement, key influencers and trending discussion themes. Access includes 25 concurrent “Buzz Monitors” (open Boolean queries across sources that measure automated sentiment across sources based on location and other parameters) and 25 Social Account monitors (monitor engagement around an owned channel such as a specific Twitter handle or Facebook page). The tool is powered by a patented human-trained algorithm and scales across the body of social data to identify patterns to measure theme, sentiment and nuanced categories. Sentiment can be annotated by the user to update sentiment algorithms. Data can be searched by location, source, user and other parameters. Content in 11 languages is monitored but there is no translation capability within the tool. Its “Live Stream” tool displays social media in near-real time in three ways – through the streaming post list, the Live Map and an hourly volume chart for the previous 24 hours. The geographic origins of posts and the relative influence of the posts’ authors register visually on the Live Map, indicating where conversation is occurring. The post list streams posts and tweets gathered by keywords, along with authors’ Klout scores. Crimson Hexagon also has a historical data library and adds a billion new posts every two days. Crimson Hexagon is not a reseller of Twitter data and acquires its data from Twitter Gnip, but all other data is available from their API. User data can also be uploaded into Crimson Hexagon for use with their analytics. Geographical data collected is based on user-reported location from user profiles, and latitude and longitude coordinates when provided. Message content is not used to identify location but is used to identify affinity groups and interests to measure affinity group size and volume of trend. In May 2014, Crimson Hexagon partnered with HootSuite Enterprise to provide Live Stream posts directly to HootSuite Enterprise.

| | |
|-----------------|---|
| Crimson Hexagon | The Basic Annual Package costs \$48,600 per year for two build users (may create collection strategies) and up to 10 basic users (may perform sub-queries on collected sets) plus one year of historical data. The API costs \$10,000 annually and is updated as features are added to the platform. API may be used to export analyzed data (e.g., authors, blog posts, sentiment, trend and volume), except raw Twitter data training, support and unlimited mentions are included. |
|-----------------|---|

B.3 DataSift

DataSift has more than 20 social media sources, including 100 percent of Twitter. Other data sources include Bitly, Facebook, Instagram, NewsCred, WordPress, YouTube, Tumblr and LinkedIn. It also supports web scraping and more than 40 analytics including natural language processing (NLP), sentiment, links, demographics-gender, topics, entity and Klout. DataSift supports 140 languages using Google Translate. Geo-location uses tweets with latitude and longitude, as well as self-identified fields. Users can define a 25-point polygon to specify an area to obtain data.

One unified API is available for historic, near real-time data across all data types, and data can be pushed or pulled, as well as raw or normalized JavaScript Object Notation (JSON). Time stamps are in GMT. A free developer trial is available at dsft.ly/dev-trial. In December 2013, DataSift announced the introduction of a new rules engine (VEDO) that features a tagging component for characterizing data. Enterprise level subscriptions start at \$5,000 per month plus data costs (e.g., Twitter 10 cents/tweet; Weibo 20 cents/1,000 posts).

| | |
|----------|--|
| DataSift | Enterprise subscriptions cost \$5,000 to \$38,000 per month. The Opinion Analysis platform costs \$38,000 per year and concurrently runs 25 different opinion analyses on one year of historical data. Depending on the subscription chosen, additional years of historical data may cost \$10,000 per year. |
|----------|--|

B.4 Geofeedia

Geofeedia is a geospatial social media monitoring company founded in 2011.¹⁹ Using location data from geo-enabled smartphones and digital photos with imbedded coordinates, Geofeedia maps tweets, photos and video posted from mobile devices within seconds. The company claims that “Geofeedia’s data set contains the hidden 70+ percent of data coming from locations compared to traditional tools that rely on certain words to be included in their data sets.”²⁰

Geofeedia tracks posts from Twitter, Instagram, YouTube, Flickr and Picasa, placing them on Bing maps down to the city block level. Users can search for a specific address, city, zip code or place name, and the results can be refined by zooming or scrolling to a specific location or by drawing a boundary around a location. Law enforcement and government entities comprise Geofeedia’s main users.

¹⁹ Geofeedia.com. <http://geofeedia.com/>.

²⁰ Geofeedia FAQ. <http://geofeedia.com/resources/faq>.

| | |
|-----------|---|
| Geofeedia | Geofeedia tracks posts from Twitter, Instagram, YouTube, Flickr and Picasa and places them on Bing maps down to the city block level. It offers a variety of social media monitoring packages. Pricing is available upon request. |
|-----------|---|

B.5 Graphika

Graphika, formerly Morningside Analytics, has raw data in Excel format that can be visualized in Graphika, allowing for improved exploratory capability for the user. The interface allows the user to drill into key conversations, power scores and most used terms. Graphika shows each segment of the map, that segment's fraction of the whole, the segment member size and power. A segment can have less members but more power, depending on who is in that segment. The user can drill into the overarching segments and look at the same metrics for sub-segments (fraction, member size and power). Graphika also visualizes several new parameters on which data can also be sorted: *Power Score* – a measure of the network density of a segment; *MScore* – a gauge of relative influence of a particular item (e.g., a Twitter account, photo) within a segment; and *Peakedness* – a measure of how quickly a concept jumps to the top of a conversation, how long it stays (displayed as a spiky or smooth graph alongside an account or content item). The user can also explore segments or sub-segments and see a lot of data. These include: key influencers and an associated MScore; current conversation leaders; key websites; key content; key tweets, photos, videos and peak dates; and latest conversation terms.

| | |
|----------|------------------------------------|
| Graphika | Pricing is available upon request. |
|----------|------------------------------------|

B.6 Hashtagify.me

Hashtagify.me displays hashtags associated with topics of interest when users enter search terms or other hashtags of interest in a “map” depicting their search terms with linkages to the 10 most closely related terms present in the same tweets. It uses the one percent of tweets that Twitter distributes for free. Languages of tweets are identified, as are percentages of occurrences of selected hashtags in those languages and top influencers.

| | |
|---------------|---|
| Hashtagify.me | Hashtagify.me offers several plans for alerting capability: personal (\$9.90 per month); business (\$69 per month); and enterprise (\$299 per month) for varying levels of services (e.g., trend analysis, top influencers/hashtag). All plans have 14-day free trial. Service offered through Hashtag.me partner Cybranding. |
|---------------|---|

B.7 HootSuite

HootSuite is based in Vancouver, Canada, and funded by venture capital. Hosted on the cloud, the HootSuite dashboard allows individuals and enterprise teams to manage their social media accounts. HootSuite has a three-tiered pricing model. The free version provides single users with five social profiles, basic reports, basic message scheduling, basic App integrators and two Rich Site Summary (RSS) feeds. The Pro version provides users with 50 social profiles, one report, advanced message scheduling, basic App integrators, unlimited RSS feeds, security, 100 archived messages and technical support for two users. At the Enterprise level, up to 500,000 users can access unlimited social profiles and reports, in addition to advanced message scheduling, unlimited App integrators, advanced security, unlimited message archiving, technical support and geotargeting. The Pro version starts at \$8.99 per month. In January 2014, HootSuite acquired uberVU, which enhances HootSuite with sentiment and influence analysis capabilities. In May 2014, HootSuite partnered with Crimson Hexagon on an integration that provides HootSuite customers posts from the Crimson Hexagon's ForSight platform streamed directly into HootSuite Enterprise version.

| | | | |
|-----------|---|---|---|
| HootSuite | The free version provides single users with five social profiles, basic reports, basic message scheduling, basic App integrators and two RSS feeds. | The Pro version provides users with 50 social profiles, one report, advanced message scheduling, basic App integrators, unlimited RSS feeds, security, 100 archived messages and technical support for two users. Prices start at \$8.99 per month. | At the Enterprise level, up to 500,000 users can access unlimited social profiles and reports, in addition to advanced message scheduling, unlimited App integrators, advanced security, unlimited message archiving, technical support and geotargeting. |
|-----------|---|---|---|

B.8 HumanGeo

HumanGeo is a New York, New York-based company founded in 2011. HumanGeo is primarily a services-based company, with a focus on using technology to meet the security and intelligence needs of clients. The company has a history of working with the federal government to map data and provide threat detection services based on social media data. HumanGeo's staff includes former technologists, operators and analysts with experience working with the U.S. Special Operations Command and Intelligence Agencies, and providing near real-time support during missions.

- HumanGeo's Media Monitor service focuses on the use of location-based smart systems to zero in on locations from which discussion trends and news are emanating. It also maintains an historical database called the GeoIndex, a collection of more than 100 million geo-located buildings and places of interest worldwide, which helps put information into useful context. Media Monitor provides clients with information and public sentiment on topics of interest through monitoring of major social media data sources including Facebook, Twitter (fire hose), Flickr, Reddit and YouTube. The

service also captures data from blogs and global social media sites generated from locations worldwide. Clients have the option of storing data on their own servers or servers operated and maintained by HumanGeo.

| | |
|------------------------------|---|
| HumanGeo Media Monitor | <p>The tool provides geo-location of news and discussion trends, maintaining a collection of more than 100 million geo-located places of interest worldwide. It provides sentiment on major social media tool messaging from Facebook, Twitter, YouTube and others.</p> <p>Pricing is available upon request.</p> |
|------------------------------|---|

B.9 InTTensity

InTTensity's flagship social media product is its Social Media Command Center (SMCC), a web-based application that looks at 75 million social media sources on an ongoing basis, including the full Twitter fire hose, Facebook and a variety of blog pages for sentiment detection. The SMCC was piloted in 2013 with the U.S. Department of State for use in sentiment detection of what citizens want and believe. Data is injected into a near real-time ingestion and orchestration engine that allows the creation of distinct processing pipelines to filter and perform specific NLP based enrichment of the social media data of interest. InTTensity does not archive the Twitter fire hose; users may save any tweets starting from their initial subscription.

| | |
|------------|--|
| InTTensity | <p>InTTensity is a suite of COTS and software-as-a-service applications including the SMCC, a web-based application that looks at 75 million social media sources on an ongoing basis (last month only), including the full Twitter fire hose, Facebook and a variety of blog pages.</p> <p>Pricing is available upon request.</p> |
|------------|--|

B.10 LM Wisdom

The Web Information Spread Data Operations Module (LM Wisdom) is a predictive analytics and big data technology tool manufactured by Lockheed Martin that monitors and analyzes open source intelligence data. The tool can ingest RSS feeds, blogs, video content, Twitter data and other social media (not identified) to create "actionable intelligence for customers." It collects, analyzes and stores structured, semi-structured and unstructured data. Using its high performance analytic algorithm, LM Wisdom analyzes content in near-real time to capture cultural context, trends, sentiment and influence. The product website advertises a flexible, user-friendly interface with multidimensional views of data in a variety of graphical and statistical outputs. Multiple languages are covered including Arabic and Farsi. LM Wisdom supports a range of scalable delivery and pricing options. The product can be accessed via secure web connection or installed

on customer specified hardware. Operations and analyst support are provided, as well as user training and assistance in the creation of customer-specific taxonomies.

| | |
|------------------------------|--|
| Lockheed Martin Wisdom | LM Wisdom is a predictive analytics and big data technology tool that monitors and analyzes open source intelligence data in near-real time. The tool can ingest RSS feeds, blogs, video content, Twitter data and other social media. It supports a range of scalable delivery and pricing options. Pricing is available upon request. |
|------------------------------|--|

B.11 Marketwired Sysomos

Sysomos offers two products – Media Analysis Platform (MAP), a dashboard that analyzes social media conversations, and Heartbeat, a near real-time monitoring and measurement tool that provides updated snapshots of social media conversations.

MAP is an in-depth research tool, while Heartbeat is a cost-effective tool designed for day-to-day monitoring and measurement requirements. MAP helps users understand which data to monitor through its full search capability. Heartbeat instances are set up for each search strategy. Sysomos provides partners with various APIs (i.e., data API, charts, trends and sentiment API, and engagement and workflow API).

Sysomos MAP employs Unicode and crawls content in 190 countries and 186 different languages. The tool covers content from 50 million active blogs, 12-13 million forums, 55,000 traditional media sources, and video coverage of YouTube, DailyMotion, Vimeo and MySpace. It also offers one year of the full Twitter fire hose. Available from Facebook are public status messages posted on user walls, personal profile pages and discussion forums from top public pages. “Authority” scoring, an indicator of how engaged and influential a source is within a given media channel is assigned to messaging on a 1-10 scale. User location is provided and is based upon user disclosed information in their platform profiles. This means that the author or source must explicitly state their location for MAP to collect location information. Collection of demographic data such as age and gender is based on user disclosed information as well as on various intelligent data extraction techniques (e.g., scanning profiles and user names for gender specific clues).

| | | |
|------------------------|---|--|
| Marketwired Sysomos | MAP is a dashboard that analyzes social media conversations including ability to search a one-year data archive. MAP is Sysomos’ full environment and billed as an in-depth research tool. ~\$2,200 per MAP seat per month | Heartbeat is a near real-time monitoring and measurement tool that provides updated snapshots of social media conversations. Working off a limited data set, it is billed as a “cost effective” tool designed for day-to-day monitoring and measurement requirements. ~\$1,000 per Heartbeat instance per month |
|------------------------|---|--|

B.12 Meltwater Buzz

Meltwater Buzz is a social media monitoring tool, primarily supporting companies as they try to understand what is being said about their products and engage with their customer base. Its web-based platform uses proprietary crawler software to monitor more than 280 million social media sites that include 100 percent of the Twitter fire hose, Facebook, blogs, forums, message boards, wiki pages and video sites. Ongoing conversations may be tracked for immediate engagement and historical discussions researched for six months prior. Meltwater Buzz is praised for its good customer service. Meltwater is a Norwegian wholly owned subsidiary of Meltwater Holdings, Inc., a company registered in the Netherlands.

| | |
|----------------|---|
| Meltwater Buzz | <p>The tool provides social media monitoring for corporate brand management and customer engagement using a proprietary crawler.</p> <p>One reference to price cited an annual license fee of \$10,275.</p> |
|----------------|---|

B.13 NetVibes

NetVibes is a personalized dashboard publishing platform composed of widgets that are pulled from a widget list open to third party developers. Common uses include brand monitoring, e-reputation management, product marketing, community portals and personalized workspaces.

The tool workspace is organized into tabs, each tab containing user-defined modules. Pages can be personalized through the use of existing themes or by creating personal themes. Customized tabs, feeds and modules can be shared individually with others or via the NetVibes Ecosystem. The interface is clean and customizable. Trade press has reviewed its RSS reader favorably (BlinkList, Laptop Magazine, 2013).

The NetVibes blog offers online help, documentation, user guide, tutorials, service upgrade announcements and maintenance schedules. Founded in 2005, NetVibes is a French company.

| | |
|----------|---|
| NetVibes | <p>NetVibes offers social media monitoring for brand monitoring, brand sentiment and e-reputation management.</p> <p>Available packages include NetVibes Basic (free), NetVibes VIP (\$3.50 per month), NetVibes Individual (\$649 per month) and NetVibes Team (pricing available upon request).</p> |
|----------|---|

B.14 Nuvi

Nuvi offers three near real-time visualizations for social media data: the bubble stream, map tool and word burst. Nuvi uses the entire Twitter fire hose and accesses the full Twitter historical archive. The tool tracks tweets in 24 languages, locations of tweets, trending topics, the number of tweets per language and trending negative terms. Public Facebook posts are accessed via

“written connectors.” Customers can create their own search strategies; more sophisticated Nuvi users will create strategies for other customers needing assistance. The Nuvi software panel is leased to client desktops for customers to log in, search and create their own set-ups.

The Nuvi overview dashboard provides statistics – volume, shared mentions, dates of spikes – which can be more deeply examined to find influencers, accounts with most reach (potential audience for a particular topic) and most spread. A topic comparison dashboard shows which topics generate more or less conversation within geographical areas.

| | |
|------|---|
| Nuvi | <p>Nuvi offers near real-time monitoring and visualization of social media tools.</p> <p>Nuvi offers product trials for 2-3 days and fee-based 30-day trials. Nuvi offers the most recent 30 days of historical data for free; beyond this timeframe, a surcharge of \$10 per day from Twitter is applied. Licensing fees vary depending on the volume of data needed by the customer. Nuvi also has partnerships with Gnip and DataSift.</p> |
|------|---|

B.15 Recorded Future

- Recorded Future’s Enterprise Web Intelligence Platform can be used on the web or brought “in house,” but is focused on “predictive analytic tools.” Using what the company calls a “temporal analytics engine,” Recorded Future provides forecasting and analysis tools to help analysts predict future events by scanning sources on the Internet and visualizing the information to show networks and patterns in the past, present and future. The tool can be used to identify events slated to occur in the future on a given date, “one month from now” and “two weeks from today.” NLP is used to determine events (100) and entities (40). The tools currently have 300,000 API-available sources, including RSS feeds, WordPress, Facebook and 1 million tweets daily. Seven languages (Arabic, Chinese, English, Farsi, French, Russian and Spanish) are supported. Geo-location of tweets is not available.

| | |
|-----------------|--|
| Recorded Future | <p>User access to Enterprise Web Intelligence Platform, forecasting and analysis tools help analysts predict future events. It currently works with 300,000 API-available sources.</p> <p>Multiple packages are available; ~\$125,000 per year for five seats, plus data API costs (~\$9,000 per month).</p> |
|-----------------|--|

B.16 SAS

SAS Social Media Analytics (SMA) tool, introduced in 2010, is a multi-module platform for monitoring social media coverage of user-designated topics. Covering more than 20,000 social media tools (e.g., blogs, Facebook, Flickr, microblogs, YouTube), users can monitor real-time Twitter traffic and generate key metrics on traffic volume, positive, neutral and negative sentiment, key influencers and word clouds. The tool can be used to look at retrospective data (unclear how much historical data are available) and create forecasts. Data can be examined at

the summary and individual document levels and may be exported in multiple standard formats. The tool supports multiple (~25) foreign languages. SAS SMA offers an iPad app.

| | |
|-------------------------------------|--|
| SAS Social Media Analytics | <p>The SMA tool offers multiple visualization capabilities for social media sources including sentiment analysis.</p> <p>Prices range from \$5,000-\$15,000 per month, plus a one-time start-up fee.</p> |
|-------------------------------------|--|

B.17 **Salesforce Radian6**

Based in California, Salesforce Radian6 is a fee-based, near-real time, online and social media feed monitor that identifies and tracks gender, location, education levels, career and lifestyle interests of Twitter message posters. Radian6 is the “social listening” component of the three-part Salesforce Marketing Cloud in which users build queries and save them in “topic profiles” to search more than 150 million sites and sources for results matching profile search criteria. The tool is advertised as a way for companies to monitor sentiment and discussion of their brand and reputation.

Five widgets typically analyze the retrieved data. These are: topic analysis – total number of posts based on a given topic profile; conversation cloud – snapshot of conversations in a topic profile; river of news – the text of tweets retrieved in a profile; topic trends – how topics trend over time; and the influence bureau – display of accounts with most vocal influence on a given topic. Filtering options include multiple foreign languages, specific sources (e.g., Twitter, Facebook, MySpace), content types (e.g., images, videos, news, blogs), and keywords which can be entered in simple or advanced modes using Boolean operators (AND, OR and NOT). Searches initially cover the last 30 days and subsequently add one day at a time over the lifetime of the topic.

A NLP positive or negative sentiment analysis capability is available through Radian6 Insights, which is a series of add-on analytical tools in multiple languages. Only two of these (Radian6 Insights and Basic Demographics) are free to add; others are fee-based. A social scoring system to measure customer abilities to influence action online and topic longevity is also available. Insights can be chosen and added to the Analytics dashboard.

| | |
|--------------------|---|
| Salesforce Radian6 | <p>The starter package for an agency includes one topic profile and includes 1 million mentions, unlimited user licenses across the full platform, an analysis dashboard, engagement console and summary dashboard for \$950 per month.</p> <ul style="list-style-type: none"> • Tier 1: Up to 10,000 hits per month (one Topic Profile (max) and five users) \$7,800 per year • Tier 2: Up to 20,000 hits per month (up to 100 Topic Profiles and 1,000 users) \$10,080 per year (bundled into contract) • Tier 3: Up to 250,000 hits per month (up to 100 Topic Profiles and 1,000 users) \$31,320 per year <ul style="list-style-type: none"> ○ Additional 50,000 hits per month \$10,440 per year ○ Additional 1,000,000 hits per month \$52,200 per year ○ Additional Dashboard User (for 10,000 Tier) \$1,200 per year ○ Additional year of historical data \$1,200 for each year of data |
|--------------------|---|

B.18 SocialMention

SocialMention is a social media search platform that culls content from social sites that rely on user-generated content, such as blogs, Facebook, Twitter, LinkedIn, Flickr, YouTube, plus over 80 other sites. Keywords and social media sites can be searched to curate a dashboard showing a variety of information regarding comments on searched terms. Terms are scored on four factors: strength (frequency), sentiment (ratio of positive to negative mentions), passion (likelihood that people mentioning the term will use it more than once), and reach (unique number of users mentioning the term divided by the total number of mentions found).

| | |
|----------------|---|
| Social Mention | <p>This social media platform search platform covers major social media tools (e.g., Facebook, Twitter). It scores results on multiple criteria (e.g., sentiment, frequency of terms).</p> <p>SocialMention is free for up to 100 queries per day. For additional queries, SocialMention charges a fee based on use, beginning at \$100 per month for up to 25,000 queries.</p> |
|----------------|---|

B.19 SocialOomph.com

SocialOomph is a tweet and post scheduler which allows free scheduling for up to five Twitter accounts and unlimited Facebook accounts. Sources include Twitter, Facebook (profiles, pages, and groups), LinkedIn (profiles, groups, and company pages), RSS feeds, blogs, Plurk, App.net and Onlywire.

SocialOomph aggregates, indexes and makes searchable a variety of information (e.g., social platforms, cataloging programs, files, tasks, email, social media and locally stored files) on one dashboard. A web-based tool, it supports some drag-and-drop capabilities across different online platforms. Platforms include Facebook, Twitter, LinkedIn, Gmail, Flickr, Picassa and Hotmail.

SocialOomph offer two versions – Free and Professional. The free version includes unlimited accounts, tweet scheduling, keyword tracking and URL shortening capability. The Professional version includes LinkedIn and Facebook scheduling, import of RSS feeds, scheduling blog posts, bulk tweet upload, delegation of account management to other users and advanced search tools. Users can access SocialOomph Twitter Unlimited Free with the option to upgrade to the fee-based account at \$35.94 per month. Users can sample a free seven-day trial of the Professional version; it does not offer API for data downloads.

| | |
|-------------|---|
| SocialOomph | <p>SocialOomph is a tweet and post scheduler.</p> <p>Users can access SocialOomph Twitter Unlimited Free (with the option to upgrade to fee-based accounts at \$6.97 for Twitter Unlimited and \$17.97 biweekly for SocialOomph Professional). A free seven-day trial of the Professional version is available with the option of staying indefinitely with the free version.</p> |
|-------------|---|

B.20 Spiral16

Spiral16 specializes in social media monitoring, sentiment analysis and business intelligence. It covers all social media outlets that are both publicly available and indexed, including blogs, forums and associated comments, as well as Facebook and Twitter. Through a partnership with Gnip, Spiral16 provides access to 100 percent of the Twitter fire hose. Geo-location is included as metadata by some data sources and is based on self-reported information only.

| | |
|----------|--|
| Spiral16 | <p>Working with Spiral16 analysts, customers create metrics and frequency of reporting to track their respective brands and topics of interest for dashboard display and custom report generation.</p> <p>Platform access starts at \$1,000 per month for up to two users.</p> |
|----------|--|

B.21 SproutSocial

SproutSocial is a fee-based social media management tool to gauge the influence of posted messages and identify demographics of followers, aggregating the social feeds on one scrolling screen. Users can generate analytical reports on engagement, shares and profile users. A web app, iOS and Android versions are available.

The tool is consistently well reviewed for its analytic capabilities, ease of use and interface; updates are easy to perform and the dashboard provides easy to use snapshots of all activities. Twitter accounts can be monitored and users' tweets can be scheduled and saved. This tool is not as mature as HootSuite; SproutSocial has been intermittently criticized in trade press for software bugs.

| | |
|---------------|---|
| Sprout Social | <p>This social media management tool is used to monitor social media channel traffic (e.g., Facebook, Google+, LinkedIn and Twitter).</p> <p>The pricing model includes a Deluxe version (\$59 per user per month), Premium version (\$99 per user per month) and Team version (\$500 per three users per month).</p> |
|---------------|---|

B.22 TweetDeck

TweetDeck is a free Twitter dashboard application that allows users to manage multiple Twitter accounts from different computers. TweetDeck by Twitter interfaces with the Twitter API for users to send and receive tweets and view profiles. Columns can be defined as home, search, interaction (chat), followers, @ mentions (“addressed to” for single or all accounts), messages (for single or all accounts), activity, users, favorites, trending, lists and scheduling.

Users can arrange an unlimited number of feeds into customizable columns in the user interface, filter feeds, schedule tweets to be sent and set alerts when new tweets of interest appear. After defining the column, the user identifies the feeds for the column. URLs can be shortened quickly. Originally developed in 2008, Twitter acquired TweetDeck in May 2011 and launched as TweetDeck by Twitter in December 2011. Trade press has praised the capabilities of the Twitter dashboard for a simple and easy-to-use interface.

| | |
|-----------|--|
| TweetDeck | <p>User logs onto a dashboard to which multiple Twitter accounts can be linked to monitor and schedule posts for brand management.</p> <p>Free</p> |
|-----------|--|

B.23 Twitter Gnip App

Gnip is a social media aggregation company providing data from many social media websites via robust streaming APIs. Gnip sources include Disqus, StockTwits, Foursquare, Tumblr, Twitter, WordPress, Facebook, NewsGator, Vimeo, Flickr, YouTube, IntenseDebate, Instagram, Google+ and Delicious. Gnip supports JSON raw format faster (quicker, as it allows stripping of formatting) and JSON Activity Stream.

Gnip Twitter packages include:

- *Historical PowerTrack* to access the archive of Twitter data back to 2006 (available either in batches or via subscription);
- *Decahose* for a statistically valid sample of at least 10 percent of all Tweets, selected at random and streamed in real time; and
- *PowerTrack* for complete coverage of user-defined filtered data.

Gnip performs near real-time enrichment (e.g., unwinding URLs, language detection, Klout scores, Klout topics and geo-location of tweets where latitude/longitude data (2 percent) or self-identified location data (30 percent) are available). Gnip testing reports testing 80 percent accuracy in location identification.

- Twitter Gnip has built an internal app on top of its Search API product that provides a sample interface for potential use by Gnip customers who may want to use it to leverage the Search API in their own products. This is a single, one-page search interface with CoffeeScript and SASS source map debugging support that allows the user to search, using simple Boolean logic, the last 30 days of Twitter. It includes a free map that shows up to 100 geo-located tweets, as well as a scrolling interface which provides the actual tweet contents. Gnip is a lightweight, inexpensive capability that could be used in the near term by organizations that need to leverage the Twitter API but do not have the expertise or funding to develop or purchase tools.

| | |
|------------------------|---|
| Twitter Gnip App | <p>Gnip is a lightweight, inexpensive user interface that leverages Twitter Search API to display search results on a map and as a scrolling list of tweets.</p> <p>The cost of \$1,750 per month includes 10,000 API requests per month. Gnip's Profile Geo Enrichment feature may be added at a cost of \$300 per month. When Search API requests exceed contractual limit, \$100 per 1,000 requests is billed monthly.</p> |
|------------------------|---|

Appendix C Research Tools

Advanced analytic research tools that are currently available at MITRE are listed below with the operational question the tool strives to answer. The following sections describe each tool in more detail.

Table 1: Current Research Tools

| Tool | Operational Question Addressed |
|--------------------------|---|
| Sentimedir | Does the public feel positive or negative about the event? |
| Tweet Explorer | How are levels of anger and other emotions changing? |
| Breakpoints | Is public opinion changing and entering some new phase? |
| Facebook Comments | What topics are being discussed in Facebook? |
| AuthorDNA | Can we determine anything about the author? Gender? Location? |
| CoFi | What topics are being discussed? Who seem to be key influencers? |
| MemeME | What ideas, media, are spreading? Which tweets are part of the meme? |
| Pinocchio | Are any groups or organizations manipulating the social media environment? |
| Event Detector | What events are emerging based on the NSWC ²¹ Dahlgren classifier? |
| Option Analysis | Can we use indicators in models to support evaluation courses of action? |
| Explore Futures | Can we run models many times to create plausible future space? |
| Visualize Futures | Can we visualize and explore the future space? |

C.1 Sentimedir

Sentimedir is a MITRE tool that detects and classifies the expression of sentiment in news, blogs, forums and tweets. Sentimedir applies statistical models to detect when a sequence of words is being used to express a negative or positive sentiment, either by the author or by someone else mentioned in the text.

The standard configuration of this metric in Sentimedir is to compute this for each day based on the local one-week window of such values, which tends to provide smoothing for this timeline and help visualize upwards and downwards trends. To further help the user interpret trends, the color of each point on the timeline is coded according to whether that day's value is more than two standard deviations larger (blue for more positive sentiment) or smaller (red for more negative sentiment) than the preceding seven days. These are displayed in the timeline below (Figure 20).

²¹ Naval Surface Warfare Center.

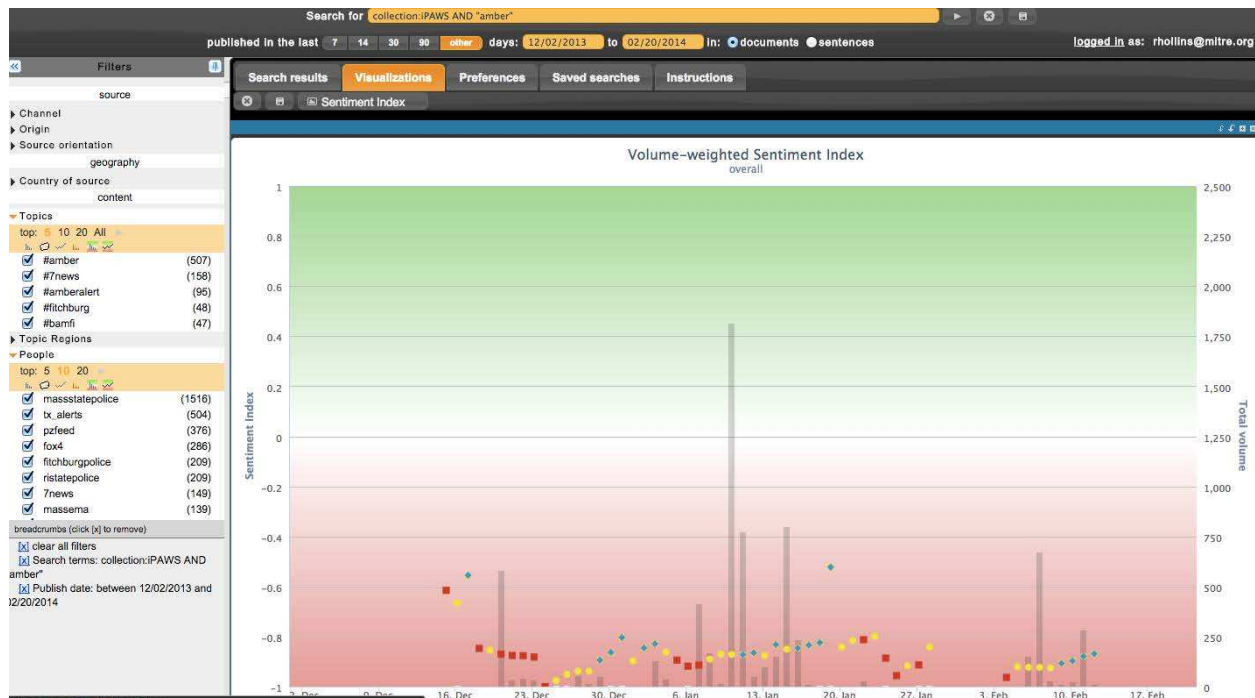


Figure 19: Sentiment Over Time

Sentimedir allows for complex searches and faceted browsing of its data, which is usually configured to access live web-hosted data at regular intervals, and then transform and index it so that the user has access to large volumes of timely information. Users can explore various aspects of the data interactively and efficiently employing features about the publication of the document (source, date, author, source location, source orientation, etc.), as well as information derived from natural language processing (NLP), such as names and descriptions referring to persons, organizations and locations, and topics identified through the application of a statistical classifier.

Examples of the topics being detected include terrorism, technology, proliferation, peacekeeping, narcotics, cultural divisions and international politics. In addition to standard document-level search, users are able to search for individual sentences that satisfy their search criteria, enabling a much more fine-grained analysis of news and blog data.

For particular installations, Sentimedir can be configured to track references to sets of person or organization names and their various aliases. Sentimedir also incorporates a range of specialized lexicons whose frequency relative to all the words can be measured and visualized. As with other measures, these can be applied to collections of documents or sentences according to their date or according to many other features of the data, as selected by the user. Some of the pre-configured word lists include those that capture references to emotions and to various domain-specific terminologies.

C.2 Tweet Explorer, Breakpoints and Facebook Comments

Tweet Explorer allows a user to explore Twitter data through three types of analysis: text analysis through Linguistic Inquiry Word Count (LIWC), temporal visualization on a timeline, and finally, by calculating breakpoints or statistically significant changes in trends.

With reference to the first type of analysis, a user can send a data set to the timeline for a specified date range and the timeline will plot volume of tweets per hour, day, week or month. In addition, the timeline can plot retweets and results from the LIWC analysis.

LIWC analysis isolates sentiment from tweets by doing a word count for different categories such as anger, anxiety or instability. LIWC calculates a ratio of word counts in a tweet to the total word count. This ratio is calculated for a minimum of 100 tweets, aggregated by a time period such as hourly, and then can be plotted on the timeline in the same way that tweet count can.

With these LIWC ratios sent to the timeline, breakpoints analysis can be done on the counts to detect statistically significant changes in phase in a set of trend lines. Users can gain greater insight into the patterns of emotion expressed as certain events unfold and understand when emotion levels are rising or falling to such a degree that greater attention is warranted. This is seen in Figure 21 with the changing colors from blue to white, where a change was detected.

A Breakpoint Analysis algorithm objectively calculates when data points (LIWC ratios, in this case) have shifted as a whole from one phase to another, as shown Figure 21.

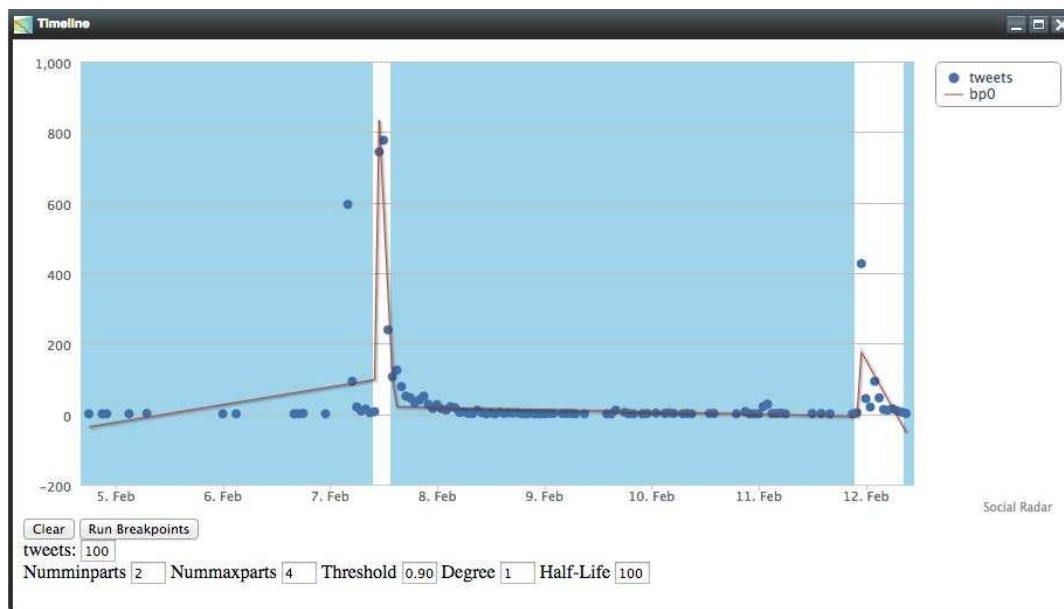


Figure 20: Breakpoint Analysis with Tweet Explorer

This algorithm has been incorporated into the Timeline widget. At each of the phase shifts, the algorithm generates a breakpoint in a graph. Between each breakpoint are phases in the data consisting of data points, and the algorithm generates a trend line that best represents these data points in each phase. The metric for deriving breakpoints is the sum of the squares of the difference between the LIWC data points and the computed trend line. The algorithm attempts to create a breakpoint between all points and to minimize the value of the metric. The number of breakpoints is also variable between 1 and 10.

Facebook Comments pulls public Facebook posts and their associated comments into Tweet Explorer for characterization through LIWC analysis. Each request sent through the Facebook Comments widget initiates a one-time pull. The 20 most recent post and their comments will appear. (The user must request posts as well as comments or will pull only comments.) Once the

comments are loaded into the Facebook Comment tool, they can be manipulated just as tweets in the Tweet Explorer are to display comments over time, or be analyzed via use of the LIWC dictionaries. To monitor a Facebook page over time, requests must be resent.

C.3 Author DNA

Author DNA (ADNA) performs classification of text to identify latent or “unseen” demographic attributes (e.g., author gender, location, native language identification) of message authors. Author attributes are drawn from message content and metadata associated with their respective social media postings. These characterizing features include the sets of words and characters used in a tweet’s text, as well as the user’s self-description, screen name, time zone, posting times, text length, emoticons, capitalization and punctuation density, and numerous other attributes. From these characterizations, a statistical profile is constructed and then compared to known samples by measuring the similarity of these distributions.

Carnie allows for users to identify Twitter authors of their location and gender, it is regarded as a cutting edge tool due to its ability to identify this information with accurate results. The tool developers have performed numerous validation experiments using Twitter data samples compiled over time. Blog profile metadata and tweet geo-location tags are used to measure the correctness of predictions. Users who produce more tweets are more likely to be tagged more accurately, so the approach is well suited for discovering highly active users. For example, 92 percent of users were placed in the correct country using tweets, measured with a vetted test set of 20,000 blog authors. Native language identification was 83 percent accurate in the identification of non-native English authors. This was determined via use of a test set of 1,000 Test of English as a Foreign Language essays representing 11 languages. Thus far, the developers have been able to label over 15 billion tweets with a country and a single computer can tag the full Twitter fire hose with gender labels in near-real time.

The approach can be used with many language and writing systems, thus is practical for many potential applications. ADNA developers’ focus has been on working with machine learning algorithms that can be trained quickly on very large amounts of training data. In this way, enormous amounts of historical data be leveraged and results applied in near-real time to large streams of tweets.

C.4 Comment Filter (CoFi)

CoFi generates clusters of Twitter data on main topics and users, and allows users to drill down to read tweets of interest as shown in Figure 22.

CoFi enables exploratory analysis of data sets in any language by using NLP techniques to group similar comments together and to prioritize messages such that the most relevant items can be identified quickly.



Figure 21: CoFi Display

CoFi makes it easier to explore, filter and understand large volumes of Twitter data by automatically discovering topics, grouping similar comments together, sorting comments by relevance and providing drilldown and timeline based visualization. The tool works with other analytics that detect elevated levels of information sharing behavior, such as retweets, and can partition social media comments by language and country of origin to provide a more targeted perspective to the analyst interested in drilling down to search messages for actionable content.

C.5 MemeME and Pinocchio

MemeME finds and displays recurring stories in large collections of tweets, grouping together messages by theme and enabling a user to quantify the reach of specific messages.

While the notion of what constitutes a single meme will vary by application, MemeME has been tuned to prefer detection of tightly defined stories. It groups together messages at a level more specific than a keyword search, but at a more general level than gathering hashtags or retweets of a particular message. For example, MemeME has grouped together the following messages as part of a single meme:

- “RT X: If Barack Obama wins the election, I volunteer as tribute for the Hunger Games.”
- “RT Y: If Mitt Romney wins the election, I volunteer as a tribute for the Hunger Games. #Obama2012.”
- “RT Z: I'd like to volunteer Mitt Romney as tribute in the Hunger Games. Thanks.”
- “If Romney becomes our president I'm volunteering myself as tribute for the next Hunger Games.”

MemeME’s algorithm can efficiently process very large data sets. The core technique performs Locality Sensitive Hashing with a back-off for exact search and has been validated on the Topic Detection & Tracking evaluation data, showing accuracy comparable to existing state-of-the-art but with an order of magnitude speedup. It has been applied to collections containing more than 100 million items. The visualization of the detected memes enables rapid understanding of a large data set while allowing drilldown to individual messages and the profiles of message authors.

Pinocchio identifies covert activity and deception in data sets by finding coordinated behaviors performed by multiple Twitter users. For example, if one person controls dozens of accounts in an attempt to simulate a digital grassroots movement (aka “astroturf”), Pinocchio will flag these accounts as suspicious.

At the center of Pinocchio is an algorithm to perform efficient search for information campaigns. Pinocchio produces an ordered ranking of suspicious actors and allows a user to investigate and flag deceptive actions occurring online. In experiments, a single user examined 544,704 messages in less than eight hours, identifying over 11 percent as suspicious or fraudulent (Figure 23).

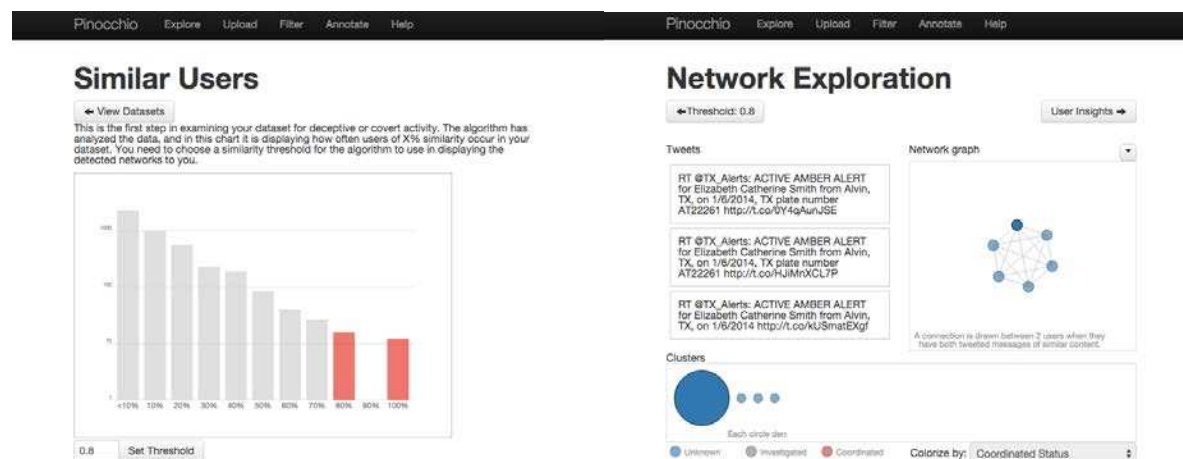


Figure 22: Pinocchio Display of Tweets with Similar Content

Unassisted, this analysis would have been highly infeasible, requiring up to 300 billion pairwise comparisons. Since these accounts are very difficult to detect through traditional means, the majority of suspicious accounts found through Pinocchio remain active on Twitter today.

The amount of deceptive activity detected by Pinocchio can vary widely depending on the source of the data (e.g., hashtag search versus randomly selected) and the goals of the analysis (e.g., spam detection versus monitoring of politically-themed activity).

Attempts to establish ground truth in a data set are similarly complicated by these factors. Pinocchio does use internal evaluation metrics, however, including the Davies-Bouldin index, to ensure that the algorithm produces tightly-connected and distinct candidate networks. Users can then examine these networks in the user interface, which supports high-level visualizations, drilldown to the level of individual messages and accounts, and the ability to annotate or flag suspicious activity for downstream filtering.

C.6 Event Detector

Event detector is designed to filter and analyze tweets in near-real time to detect statistically significant changes. It uses a Twitter/Gnip feed, which is a 10 percent pseudorandom sample of all publicly available tweets. Each tweet is geo-location-tagged using both MITRE’s ADNA algorithm, which provides a country code and the commercial Metacarta tagger shown in Figure 24.



Figure 23: Event Detector Display Showing Detected Events from Tweets

To develop an event detector, queries using keywords, hashtags, phrases, wildcards, locations and metadata to retrieve a set of tweets are needed. The tweets can also be annotated to indicate which are relevant or irrelevant. An automated algorithm that trains a support-vector machine to further filter the tweets is run. Each set of filtered tweets is then separated into geographical regions using a constrained Voronoi²² decomposition. Next, each geographical region's subset of the tweets is analyzed to determine whether a statistically significant change is taking place. The noise mean estimate will vary depending on the day of the week and the hour of the day, so a sliding median filter is applied separately for each of the 168 combinations to yield noise mean estimates that account for these variations. The noise variance is estimated using the Gamma Test (Jones et al., 1997). Finally, detections are produced by examining each span of time (whose length is set by a temporal resolution parameter) with a statistical confidence interval test to determine if any of them significantly exceed the noise baseline.

The event detector has been tested on automatic detection of earthquakes in the continental United States, FIFA Confederation Cup matches, and has also been used in the past to successfully detect alleged uses of chemical weapons in Syria, including an attack on August 21, 2013.

C.7 Options Analysis, Futures Explorer, Visualize Futures

The integrated capabilities described in this section provide the ability to compare the robustness of different courses of actions. The Options Analysis widget uses content analysis indicators with short-term forecasts in models that can assess evidence and/or evaluate the impact of one course of action under a given set of conditions. Option Analysis has several components: 1) connects to near-real time data; 2) performs a short-term forecast of that data; 3) integrates models that utilize that data (currently requires a developer to write code); and 4) runs one possible future once the model is integrated.

²² For more information: http://en.wikipedia.org/wiki/Voronoi_diagram

Futures Explorer is a complementary capability. It is an exploratory modeling engine that allows users to compare the robustness of multiple courses of action. Users select their courses of action and any intermediary variables that are a source of uncertainty. The Futures Explorer widget provides an exploratory modeling capability that runs the model many times, for multiple courses of action under a range of conditions. This allows for the robustness of each course of action to be analyzed and compared. Futures Explorer will then generate a number of different model runs for each course of action by sampling over a distribution of possible values for each intermediary variable. The user can select a normal or equal distribution, or create a more complicated distribution with a set of sliders. The end result is a set of plausible outcomes for each course of action.

The Visualize Futures widget graphically displays the data from Futures Explorer, so that users can compare the robustness of each course of action. In this way, users are provided with “option awareness.” The users view the results using the Visualize Futures widget, which represents the distribution of plausible outcomes for each course of action as a box plot. The Visualize Futures widget provides a visualization that enables a user to compare the robustness of multiple options given the data generated by Futures Explorer.

For this effort, the following model was built:

- The goal of the model is to facilitate understanding of the relationships between public sentiment and the types of Wireless Emergency Alert (WEA) messages sent to the public.
 - The model is an Agent Based Model (ABM) where agents or people who tweet are modeled as individuals in the populations.
 - An ABM is a simulation composed of at least three aspects: agents, a meaningful environment for the agents to exist in and a set of rules that decides what happens when one agent interacts with the other, or potentially when an agent interacts with the environment.
 - Interaction between agents is dictated by rules, but rules may contain random components that change the predictability of those interactions. Because of this, ABMs are great at representing human systems, since different rules can be created for different types of human entities.
 - The goal of an ABM is to simulate the actions and interactions of autonomous agents (in this case, people) and to understand the effects on the system as a whole.
 - This modeling approach was selected so individuals can be modeled to have various cell phone providers, can opt in or out of alert services, can have phones that are or are not WEA capable, and can move in and out of carrier range.
- The model combines information from the WEA messages, the cell towers to which the messages were sent, with the location of the people on Twitter who commented. The model specifically integrates numbers of persons in the general population, cell tower locations, alerts issued and population sentiment towards the alert. This allows sentiment to be mapped to locations and compared with WEA message reach by location.
 - Commercial mobile service providers (e.g., AT&T, Sprint, Verizon) are permitted to choose how to geographically distribute alerts. For example, Verizon

distributes alerts to towers in a “bounding box,” which is a polygon that generally covers an area more tightly. Other carriers distribute alerts to towers mapped to a code representing a particular county.

- To model population areas receiving cellular service, the following is needed: cell tower latitude and longitude, size of cells that each tower serves, service provider operating on the tower, and, if possible, raw observations to discern cell size in cases where the cell may not be circular. The Federal Communications Commission (FCC)²³ free data source from June 2012 was used in the model. FCC’s data contains 23,498 cell towers distributed across the U.S. Historic cases cell tower counts are: Texas = 1,791, Arizona = 473, Mississippi = 475, Hawaii = 82, Rhode Island = 28 and Massachusetts = 149. The data set does not contain cell sizes or raw observations.
- Population estimates by zip codes were obtained from Splitwise²⁴ and market share of wireless subscriptions held by carriers in the U.S. was obtained from Statista.²⁵
- To quantify sentiment in tweets pertaining to WEA alerts, a set of 396 geo-located tweets with latitude and longitude coordinates from the users’ device were obtained between March 2013 and September 2014.
 - The geo-located Twitter data was compared to weather alerts over the same time period. It is very difficult to deduce the alert to which a person is reacting in a given tweet. This was attempted using only geo-located tweets, then isolating tweets around the time of the alert. Finally, the distance of the tweet’s location to the centroid of the alert’s polygon was determined, with the assumption that the National Weather Service would arrange for the centroid to be approximately where the weather incident would occur.
 - The tweets were then partitioned by distance: within polygon, as well as 20, 40 and 80 miles from the center. The tweets were assigned to the first category for which they qualified, which led to 53, 35, 48 and 137 tweets in each category, respectively.
 - The tweets were run through Tweet Explorers LIWC algorithm for positive and negative emotion (Appendix C.2). The LIWC negative emotion output was averaged over tweets in each category and normalizing by the number of tweets in each month.
 - For a given transmission method, the cell towers within the respective areas are determined using the cell tower location and zip code data within the range of the tower. The location of the tower is compared with the center of the zip code area. A constraint is that the center of the zip code area needs to be within two miles of the tower. From this, the census population that will receive the alert is calculated. The current model assumes the entire population of the zip code receives the alert if the cell tower is in range, constrained by the number of people over 18 and

²³ <http://catalog.data.gov/dataset/fcc-geographic-information-systems/resource/9396de79-f935-423d-8370-c53564376b9c> (use file called cellular.dbf)

²⁴ 2010 US Population Census by Zip Code. <http://blog.splitwise.com/2013/09/18/the-2010-us-census-population-by-zip-code-totally-free/>

²⁵ <http://www.statista.com/statistics/199359/market-share-of-wireless-carriers-in-the-us-by-subscriptions/>

people who have a cell phone that is WEA capable. Key parameters of the cell tower population are calculated such as distance of the zip code from the centroid of the bounding box, number of alerts received, and the time it was received at the zip code.

Appendix D National Weather Service Site Visit

The goal of the study was to get a specific understanding of steps taken and tools used by a National Weather Service (NWS) Forecast Office.

D.1 NWS Alerting

Forecasts are initially made by the 122 local NWS offices in the U.S. Once the warning is created locally it is sent to NWS Headquarters where it is automatically converted to Common Alerting Protocol (CAP). The CAP message is pushed to the Integrated Public Alert and Warning System (IPAWS). If a message qualifies as a Wireless Emergency Alert (WEA) message, based on the type of message, severity, certainty and urgency of the event, it is sent from IPAWS to the cell providers who broadcast it to cell phones in the affected areas.

Figure 25 shows a tool used by NWS to monitor weather using radar, in this case a tornado. The figure shows a polygon used to highlight the region where a warning is being issued.

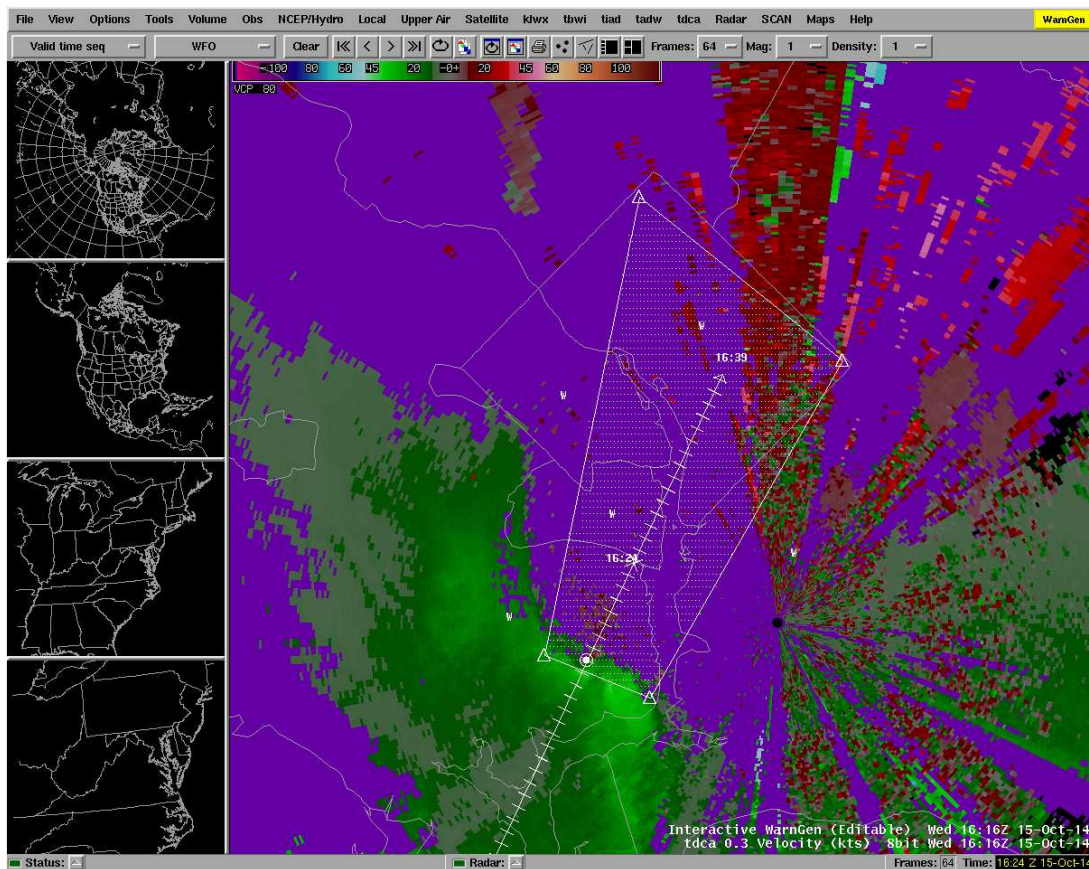


Figure 24: NWS Radar with Alert Box

Once the polygon is defined, NWS uses WarnGen²⁶ software to define messages about warnings, watches, statements and forecasts. Figure 26 shows an example of this software. In this instance, the “Tornado Warning” is highlighted.

The screenshot displays the WarnGen software interface with the following sections:

- Backup:** A button labeled "LVX".
- Track type:** Radio buttons for "One Storm" (selected), "Line of Storms", and "Box and Track".
- Edit:** Radio buttons for "Box", "Track", and "Box and Track".
- Redraw Box on Screen from:** A button labeled "Track".
- Warned Area Visible:** A checkbox that is currently unchecked.
- Warned/Hatched Area:** A button labeled "Warned/Hatched Area".
- Dam Break Threat Area:** A button labeled "Dam Break Threat Area".
- Product type:** A list of product types with "TOR - Tornado Warning" selected. Other options include "SVR - Svr Tstm Warning", "Severe Weather Statement", "Short Term Forecast", "SMW - Special Marine Warning", "Marine Wx Statement - follow SMW", "FFW - Flash Flood Warning", "Flash Flood Statement", and "Other: Areal Flood Warning". An "UPDATE LIST" button is next to the list.
- Time range:** A section with a "Duration:" dropdown set to "30 min" and a time range from "16:24 Wed 15-Oct" to "16:54 Wed 15-Oct" with a "Change..." button.
- Optional bullets:** A text area containing several bullet points, with "- Radar indicated..." highlighted. The bullets include:
 - ***** BASIS FOR WARNING (CHOOSE 1) *****
 - *****Tornado Emergency (ask Lead Forecaster)*****
 - Radar indicated...
 - Confirmed tornado...
 -by Spotters
 -by Law Enforcement
 -by County Officials
 -by Other Sources
 - History of tornadoes...
 - ***** CALLS TO ACTION (CHOOSE 1 OR MORE) *****
- Instructions:** A text area with the instructions "Adjust Centroid in any Frame" and "Adjust box around Warning Area".
- Buttons:** "Create Text", "Restart", and "Close" buttons at the bottom.

Figure 25: NWS Warning Software and Message Output

²⁶ This is a custom product for the National Oceanic and Atmospheric Administration (NOAA).

Figure 27 shows the message that is created after clicking “create text.” This is the message that is received at NWS Headquarters and converted to a CAP message for transmission to IPAWS.

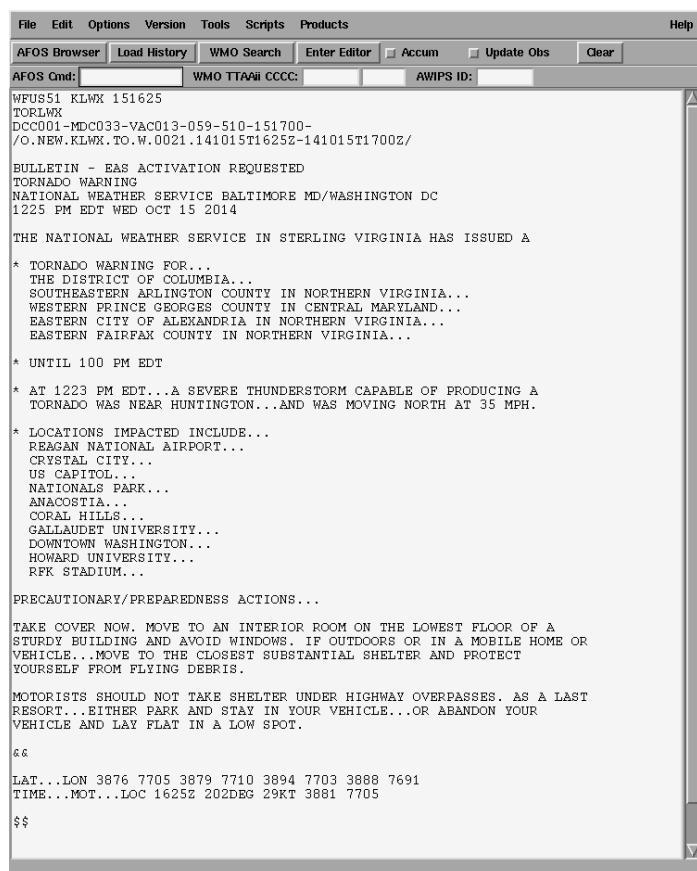


Figure 26: NWS Weather Message

After issuing an initial watch or warning, the alert may be updated, expired or cancelled. The statement may also update information, such as confirmed touchdown or damage. This information is sent using the same process outlined for the initial alert. As WEA messages are limited to 90 characters, much of the detail is not sent and is only available by other distribution media.

D.2 NWS and Social Media

NWS is interested in monitoring the public’s response to weather events as indicated by the fact that many offices are currently using tools to monitor Facebook and Twitter for weather-related keywords in online posts. For example, as it may be as many as 70 percent²⁷ of tornado warnings do not result in a touchdown and when they do occur they usually effect only at the neighborhood level, NWS wants to insure that the public understands this and still takes warning seriously and respond appropriately.

²⁷ Mike Gerber, NOAA Headquarters

One local office is currently using the free version of the commercial tool HootSuite to follow relevant weather-related keywords on Facebook and Twitter. In this office, Facebook has been monitored for 2-3 years, and Twitter has for one year. They use keywords in near-real time to watch trends related to general weather, winter weather, severe weather, flood search, etc. The amount of time spent analyzing this data varies by the Alert Originator's (AO) workload.

Figure 28 shows the HootSuite dashboard. Across the top, Facebook, Twitter, Twitter List and other tabs are shown. The lefthand column is that office's Twitter feed. The middle column shows a monitor of relevant keyword mentions. The righthand column shows the severe weather search, which is restricted geographically as denoted by the "geocode" in the thunderstorm header. HootSuite, therefore, allows the AO to monitor trends related to keyword searches in a dashboard.

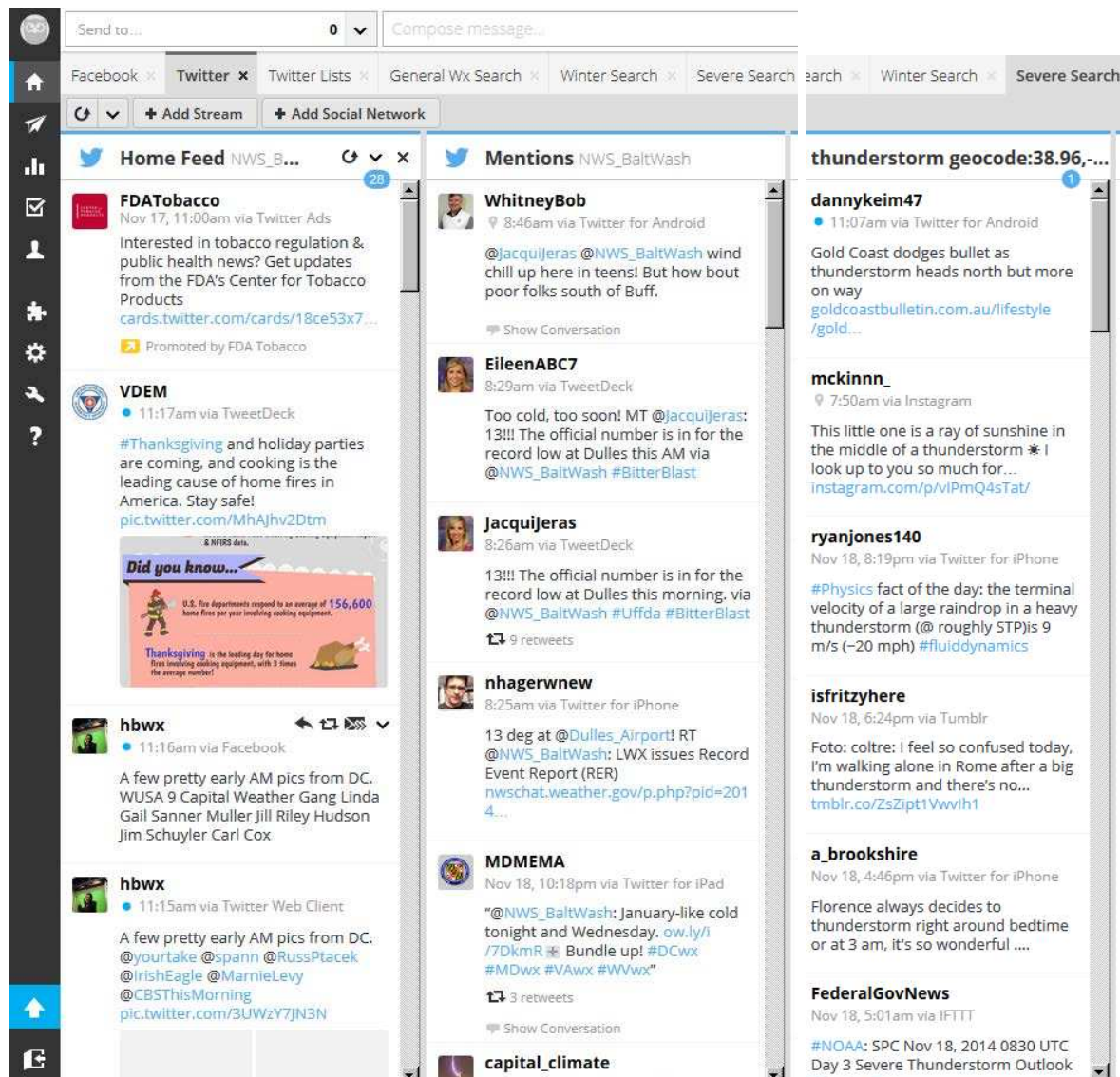


Figure 27: Example Use of HootSuite to Follow Facebook and Twitter

This office is also evaluating a Social Media Dashboard shown in Figure 29. This used searches similar to HootSuite including flood, water rescue, high water, etc. One of the benefits of this dashboard over HootSuite is that pictures are more available in the dashboard.

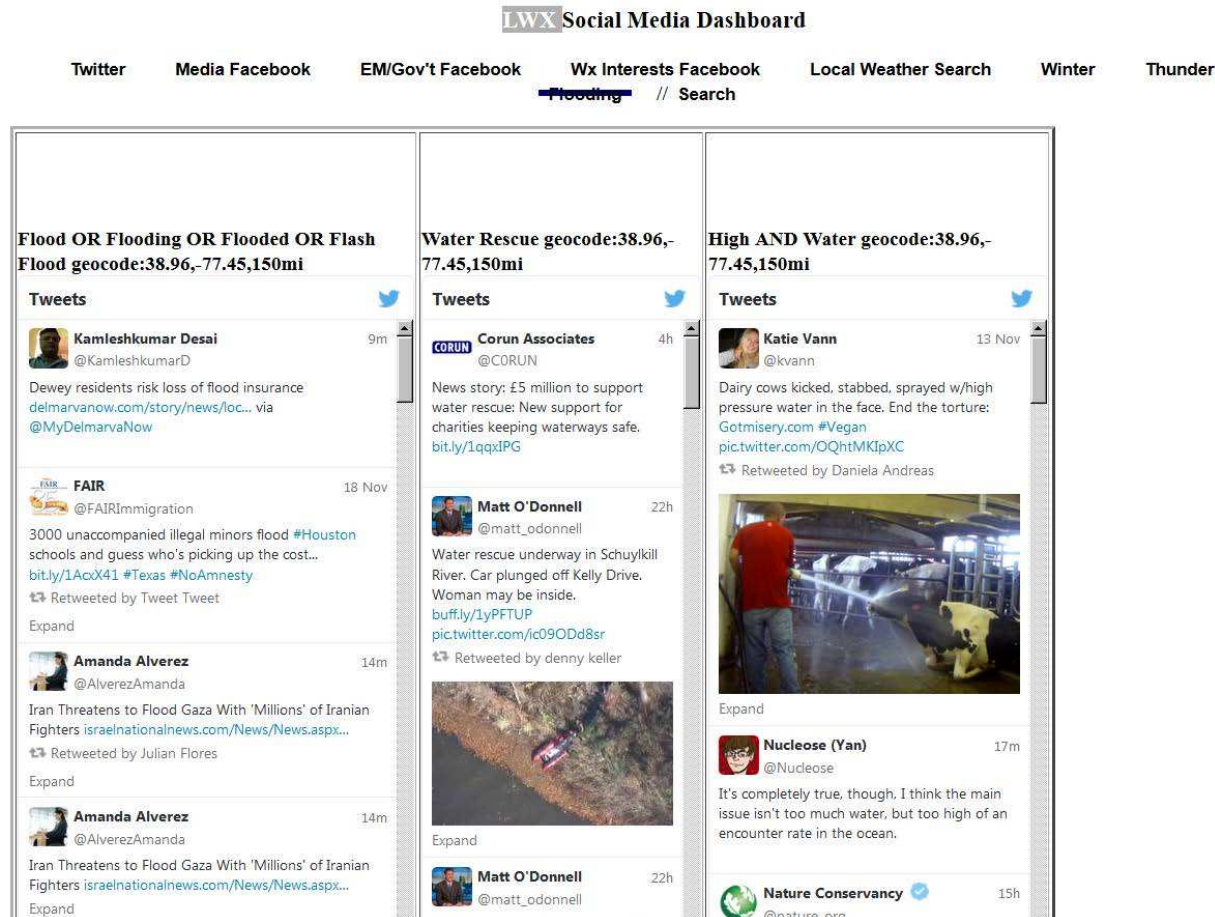


Figure 29: NWS Capability for Monitoring Twitter

Appendix E Acronyms

| Acronym Term | Acronym Definition |
|-----------------|---|
| AMBER | America's Missing: Broadcast Emergency Response |
| AO | Alert Originators |
| API | Application Programming Interface |
| CoFi | Comment Filter |
| DHS | Department of Homeland Security |
| EAS | Emergency Alert System |
| FCC | Federal Communications Commission |
| FEMA | Federal Emergency Management Agency |
| FRG | First Responders Group |
| HS HSEDI | Homeland Security Systems Engineering and Development Institute |
| IPAWS | Integrated Public Alert and Warning System |
| JSON | JavaScript Object Notation |
| LIWC | Linguistic Inquiry Word Count |
| MAP | Media Analysis Platform |
| NLP | Natural Language Processing |
| NOAA | National Oceanic and Atmospheric Administration |
| NWS | National Weather Service |
| RSS | Rich Site Summary |
| S&T | Science and Technology Directorate |
| HSEDI | Systems Engineering and Development Institute |
| WEA | Wireless Emergency Alert |
| WISDOM | Web Information Spread Data Operations Module |