Tracing and Responding to Foodborne Illness

Rebecca Tegtmeyer
Michigan State University
Kresge Art Ctr
600 Auditorium Road Room 26C
(517)432-1613
tegtmey2@msu.edu

Liza Potts
Michigan State University
WIDE Research
235 Bessey Hall
(517) 355-2400
lpotts@msu.edu

William Hart-Davidson
Michigan State University
WIDE Research
7 Olds Hall
(517) 353-9184
hartdav2@msu.edu

ABSTRACT
In this poster, we describe a how we use social web tools to track, trace, and respond to foodborne illness. Using a combination of data streams, analytical tools, bots, and dashboards, we propose solutions to the current challenges facing government officials, NGOs, and everyday people.

Categories and Subject Descriptors
H.5.2 [Information Interfaces and Presentation]: User Interfaces – *Natural language, Screen design, User-centered design*.

General Terms
Design, Experimentation, Human Factors

Keywords
Dashboards, Twitter, data mining, foodborne illness, disaster

1. INTRODUCTION
Foodborne disease outbreaks are particularly difficult to detect due to the highly complex and distributed nature of our food supply chain. In the short term, digital disease detection techniques that have been successfully used to detect flu and other infectious diseases could decrease the time it takes to detect outbreaks of foodborne illness. In the long term, digital disease detection data could be combined with sensor data collected throughout the food chain to form a "smart grid" for a more comprehensive system of rapid detection, warning and traceback.

2. ABOUT BIOSURVEILLANCE
2.1 The Problem
Each year 48 million Americans get sick, 128,000 are hospitalized and 3,000 die from foodborne illness. The current system for identifying outbreaks is highly labor-, and more importantly, time-intensive [1]. The CDC estimates that it takes on average 5-28 days to identify a case as part of an outbreak. Furthermore, for every case of salmonellosis reported, 28 cases go undiagnosed. This 28-fold attenuation in signal strength severely impairs the ability of the system to rapidly identify clusters of cases that public health officials rely upon to home in on the source of contamination [1].

2.2 The Opportunity
It has been demonstrated that digital disease detection can match or surpass traditional techniques for detecting infectious disease outbreaks such as flu and dengue [1, 2]. Early examples used data search queries (e.g. looking at how people searched for information on flu), but have progressed to include new streams of data, such as Twitter[2, 3]. In addition to passively looking at new data streams, there have been examples of users actively organizing themselves on Facebook for successful disease outbreak investigation. These techniques could be particularly useful for detecting outbreaks of foodborne illness where the vast majority of victims do not seek medical attention and, therefore, would not be detected under the current system, yet many of these individuals are likely to self-report via social networks [3].

2.3 Challenges
First, it will be necessary to demonstrate proof of concept using existing and emerging ***data streams*** (search queries, Twitter, Facebook, etc.) and ***analytical tools*** (intelligent search engines or robots that can detect an increase in volume and/or clusters in space and time). This will require developing a specific ***ontology***: verbs and nouns, their misspellings and associations that act as indicators of foodborne illness. Historic data from a recent outbreak will be procured for an initial test of sensitivity, and then the model can be improved to proactively "listen" for new outbreaks.

3. TRACING EVENTS
Tracing foodborne illness through for the sake of mitigating a disaster involves issues concerning time and space. By this we mean that analytical tools must be able to locate the appropriate data stream for a given culture, location, and timeline for a specific event. In this section, we define these concepts further.

3.1 Across Time
Foodborne illnesess are bound by fairly standard lengths of time when the first outbreak occurs and the last person falls ill. Bound by time, being able to track specific issues across specific blocks of time is essentially for early detection. Using timestamps found in social web tools is one way to locate foodborne illnesses online. For example, Twitter tweets and Facebook status updates have timestamps, as do Yelp, Foursquare, and Facebook check-ins.

3.2 Across Space
Foodborne illnesses are also typically located in clusters in space. Using location-aware information from Twitter, Facebook, Yelp, and Foursquare would greatly aid in this detection. Of course, there are outliers – such as conferences, conventions, and tourism. However, the data sets are typically large enough to show clusters.

4. RESPONDING
In order to respond to people during these outbreaks, we are using a three-pronged approach. Focusing on tracing, responding, and mitigating, we are working to deploy Twitter response robots,

dashboards for organizers and officials, and linking to content-rich web sites and tools.

4.1 Deploying Robots

Responding to specific Twitter status updates, these robots (known as "bots") will reply to whoever initiates the tweet. The tweet response will notify the user about recent issues and encourage the participant to engage further on one of the partner websites. Engagement can be simply learning more about foodborne illness or participating in surveys about a current emerging outbreak. Future work will include building response bots for Foursquare, Facebook, and other social web tools.

4.2 Launching Dashboards

The major focus of our work is on this dashboard system. The system will allow for different levels of participation from users, a curious browser, affected reported, and active reviewer.

The basic components of the dashboard include a social media stream, tag cloud, location map, timeframe graph, and a form to submit reports. The social media stream pulls keywords and clusters related to foodborne illness preset by the system. It also displays invalid, valid, and pending reports. A user can locate the posts on the map and choose to login as an active reviewer to submit the post for further validation.

The tag cloud is animated and displays keywords and clusters from the social stream. The scale of words indicates the amount of usage in posts. A curious browser can add words of relevance to the tag cloud.

The map displays possible outbreaks by showing confirmed and unconfirmed reports. A curious user can search various locations within the U.S. The active reviewer can track, find, share, and save multiple outbreaks based on location.

The timeframe graph displays confirmed and un-confirmed reports correlated with time and the location. All users can interact with the timeframe graph to display data from the past, going back 14 days.

The report submission form is for affected reporters, they can chose to submit as a guest or create a reviewer profile. In the report submission form the affected reporter provides detailed information about their case and publishes it to the system.

4.3 Linking Websites

In order to close the feedback loop, both the robots and the dashboards will provide links to more robust content. In particular, we are exploring ways in which timely information from the Centers for Disease Control (CDC) and other organizations would benefit these response mechanisms.

5. Future Work

Early detection of foodborne illness outbreaks will substantially reduce casualties and their associated costs due to medical expenses, lost productivity, and damage to the food industry from loss of consumer confidence. The emergence and global adoption of social networks holds promise as a tool for early signal detection, but its usefulness will be realized only if it is widely accepted by public health agencies responsible for food safety. By their nature, regulatory agencies are slow to adopt new technologies and this will likely be the case for a social media-based system with its attendant data quality shortcomings. It is important, therefore, to develop a well-organized adoption strategy as an integral component of a web-based food safety biosurveillance project.

6. REFERENCES

[1] Brownstein, J.S., Freifeld, C.C.,& Madoff, L.C., M.D. (2009). Digital disease detection: Harnessing the web for public health surveillance. *N Engl J Med* 360. 2153-2157

[2] Hesse, B.W., & Shneiderman, B. (2007). eHealth research from the user's perspective," *Am. J. Preventive Medicine 32*.5S. S97-S103.

[3] Eysenbach, G. (2009). Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the internet. *J. Medical Internet Research*11.1 e11.

[4] Jones, D. and Potts, L. 2010. Best practices for designing third party applications for contextually-aware tools. In *Proceedings of the 28th ACM International Conference on Design of Communication* (Sao Paolo, Brazil). SIGDOC '10. ACM, New York, NY, 95-102.

[5] Potts, L. 2009. Designing for disaster: Social software use in times of crisis. *International Journal of Sociotechnology and Knowledge Development* 1, 2, 33-46.

[6] Potts, L. 2009. Peering into disaster: Social software use from the Indian Ocean earthquake to the Mumbai bombings. In *Proceedings of the International Professional Communication Conference*. Hawaii: IEEE.

[7] Potts, L. and Jones, D. 2011. Contextualing experiences: Tracing the relationships between people and technologies in the social web. *Journal of Business and Technical Communication* 25, 3, 338-358.

[8] Potts, Seitzinger, Jones, and Harrison. 2012. Tweeting Disaster: Hashtag Constructions and Collisions. *Proceedings of the 29th ACM International Conference on Design of Communication* (Pisa, Italy). SIGDOC '11. ACM, New York, NY.

[9] Vander Wal, T. 2007. Folksonomy. *Vanderwal.net*. Available from http://vanderwal.net/folksonomy.html.

Conference'12, October 3–5, 2012, Seattle, Washington, U.S.A.