

Chapter 14

HEALTHMAP

HealthMap is a freely accessible, automated real-time system that monitors, organizes, integrates, filters, visualizes, and disseminates online information about emerging diseases. The goal of HealthMap is to deliver real-time intelligence on a broad range of emerging infectious diseases for a diverse audience, from public health officials to international travelers.

HealthMap.org Web site has been operational since September 2006 (Figure 14-1). US Health and Human Services and the US Department of Defense among other national or international organizations have used their data stream for surveillance activities. HealthMap currently receives approximately 15,000 unique visitors per month from around the world.

Figure 14-2 shows the system architecture of the HealthMap application, which consists of the following components: (1) data acquisition, (2) information characterization, (3) signal interpretation, and (4) dissemination and alerting.

The system acquires multistream data automatically every hour from a variety of electronic sources: online news wires, Really Simple Syndication (RSS) feeds, ProMED mailing lists, and EuroSurveillance and WHO alerts. The text data are automatically categorized into groups by disease types and locations with text mining techniques. The system now handles information in English, Spanish, and French (Brownstein et al., 2008a). HealthMap currently processes 133.5 disease alerts per day on average (95% Confidence Interval: 124.1–142.8), with approximately 50% categorized as Breaking News (65.3 reports/day). With a 30-day default window, the system may display over 800 Breaking News alerts on a given day (Freifeld et al., 2008).

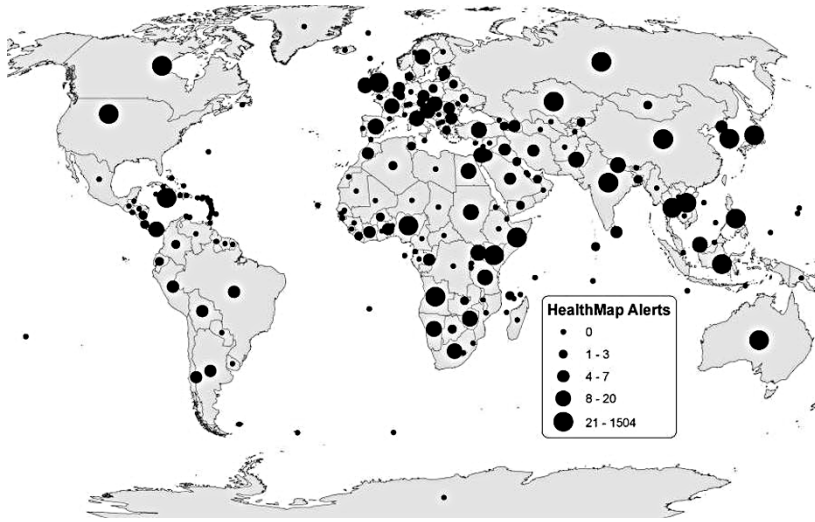


Figure 14-1. HealthMap geographic coverage, October 1, 2006 to February 16, 2007 (Freifeld et al. 2008).

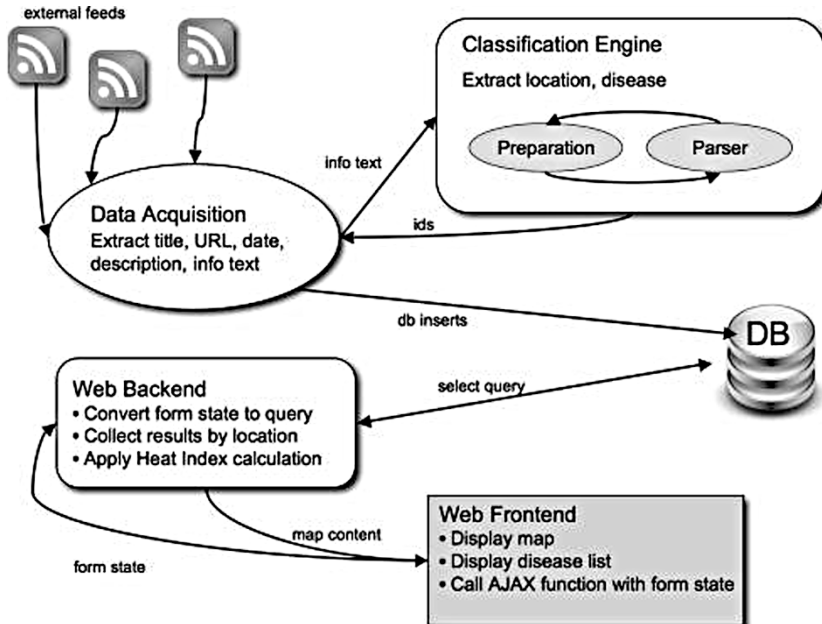


Figure 14-2. Framework for Internet-based surveillance (Freifeld et al. 2008).

HealthMap aggregates the disease reports by source, disease, and geographic location. This information characterization is performed using natural language interpretation and automated text mining and parsing techniques. The characterized information is then overlaid on an interactive map (supported by Google Maps), which allows for user access to the original report. On the left-hand panel, the HealthMap page allows improved information filtering by feeds sources, disease, and countries.

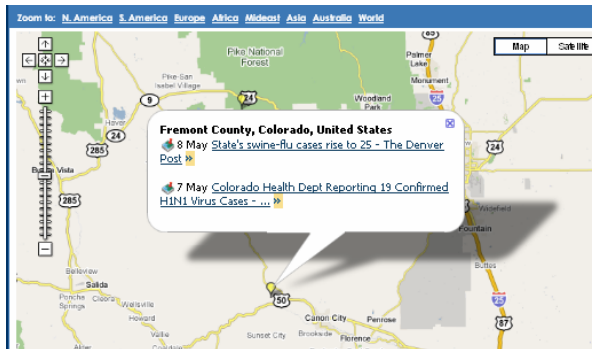


Figure 14-3. HealthMap page showing the latest information on H1N1 Flu as of May 27th, 2009 (lower-left corner: bringing up the related news at a particular location as zooming out). (source: Healthmap Web page).

In April 2009, a new strain of influenza known as swine flu (H1N1 flu) was first detected in the US and soon led to an outbreak in Mexico. It is now present in over two dozen countries around the globe including Canada, Japan and the UK. HealthMap began aggregating and filtering real-time information on the novel flu virus on April 1, weeks before the news emerged in English-language resources. HealthMap tracked early reports from the Mexican press on a “mysterious” influenza-like illness occurring in the town of La Gloria in the state of Veracruz that reportedly infected 60% of the 3,000 residents and killed 2 people.

Figure 14-3 shows a global alert map of the H1N1 disease during its 2009 outbreak as of the end of May 2009. Zooming to a specific region and clicking on a balloon bring up a list of disease related news articles at that region.

HealthMap represents a successful practice of mining the Internet for public health surveillance purposes to support and enhance the traditional public health infrastructure. It demonstrates that news reports in particular can be a valuable resource for information as inherently the media has the ability to saturate towns, cities, and communities where public health officials may or may not be present to report on potential disease outbreaks.

We provide the following project link and some key readings for the readers who might be interested in learning more details about the HealthMap Project.

Project link:

<http://www.healthmap.org>

Important readings:

1. Freifeld, C. C., et al., “HealthMap: Global Infectious Disease Monitoring through Automated Classification and Visualization of Internet Media Reports.” *Journal of the American Medical Informatics Association* 2008. 15(2): pp 150–157.
2. Chute, C. G., “Biosurveillance, Classification, and Semantic Health Technologies.” *Journal of the American Medical Informatics Association* 2008. 15(2): pp 172–173.
3. Brownstein, J. S., C. C. Freifeld, B. Y. Reis, and K. D. Mandl, “Surveillance Sans Frontières: Internet-Based Emerging Infectious Disease Intelligence and the HealthMap Project,” *PLoS Medicine* Vol. 5, No. 7. (1 July 2008), e151.