

Development of a Vietnamese Language Outbreak Mapping and Surveillance System

David Bui*^{1,2}, Sumiko Mekaru^{1,3}, Clark Freifeld^{1,6} and John S. Brownstein^{1,5,4}

¹Children's Hospital Informatics Program, Division of Emergency Medicine, Children's Hospital Boston, Boston, MA, USA; ²Dept of Epidemiology, Mel Enid Zuckerman College of Public Health, Tucson, AZ, USA; ³Dept of Epidemiology, Boston University School of Public Health, Boston, MA, USA; ⁴Dept of Pediatrics, Harvard Medical School, Boston, MA, USA; ⁵Dept of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada; ⁶Dept of Biomedical Engineering, Boston University, Boston, MA, USA

Objective

To present the development of a surveillance system utilizing on-line Vietnamese language media sources to detect disease events in Vietnam and the South East Asian Region.

Introduction

In the South East Asia Region (SEAR), infectious disease continues to be a leading cause of death. SEAR countries, like Vietnam, are also at risk for outbreaks of emerging diseases due to high population density, proximity to animals and deforestation. (1) Given Vietnam's location in SEAR and its recurrent outbreaks of zoonotic diseases—timely surveillance in Vietnam is critical to global public health. (1) Online news sources have been recognized as potential sources for early detection of emerging disease outbreaks, as was the case with SARS. (2) HealthMap, an innovative disease surveillance system developed at Boston Children's Hospital, leverages the expediency of online news media by using text-mining technology to monitor and map global disease outbreaks reported by news sources.

Methods

HealthMap currently monitors disease related news in 15 languages. Building on HealthMap's existing infrastructure, we translated the system's existing disease and location name dictionaries to corresponding Vietnamese terms to train the system to detect disease and locations cited in Vietnamese news sources. To ensure comprehensive capture of disease terms, both formal disease names and colloquial synonyms were used. Vietnamese locations were sourced through the official Vietnamese government website. Search queries were developed using a variety of outbreak related terms like “bùng phát” (outbreak) or “bệnh” (disease), and specific disease names like “Cúm gia cầm” (Avian Influenza). Automated searches are performed in the Vietnamese version of Google News.

Results

As of August 18, 2012, after 2 months in operation, the system has mapped 433 alerts in 27 diseases reported in the Vietnamese media compared to 7 diseases in English in the same time period. The collected alerts were mapped to 699 province level and district level (or lower) locations compared to only 16 in the English feed. To date, the system collected 38 alerts of Avian Influenza from Vietnamese sources compared to only 2 in English sources; 30 alerts of Dengue compared to 7 in English; and 25 alerts of Hand Foot and Mouth compared to 6. The system also collects outbreak case counts in Vietnam. For example, counts of human Dengue cases in 97 locations in Vietnam were collected, providing a rich dataset for monitoring epidemic spread and progression. The surveillance feed also received 2 reports of outbreaks in crops. Zoonotic disease outbreaks in Vietnam were more comprehensively covered in the Vietnamese feed compared to English.

Conclusions

Leveraging the expediency of freely available online news media, the developed surveillance system is able to detect and map outbreaks occurring in Vietnam in near real-time, providing health organizations and researchers with timely and comprehensive coverage of disease events to assess pandemic risk and mobilize aid as necessary.



Screenshot of HealthMap Vietnamese Alerts

Keywords

Surveillance; Infectious Disease; Vietnam; HealthMap; Mapping

Acknowledgments

Sumiko Mekaru, John Brownstein, Clark Freifeld, Susan Aman, and the whole HealthMap gang.

References

1. World Health Organization. Combating Emerging Infectious Diseases. http://www.searo.who.int/LinkFiles/Avian_Flu_combating_emerging_diseases.pdf. Accessed August 28, 2012.
2. J.S. Brownstein. Digital Disease Detection: Harnessing the Web for Public Health Surveillance. *N Engl J Med* 2009; 360:2153-2157 Available at: <http://www.nejm.org/doi/full/10.1056/NEJMp0900702>. Accessed April 16, 2012.

*David Bui

E-mail: davidbui@email.arizona.edu



ISDS Annual Conference Proceedings 2012. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 3.0 Unported License (<http://creativecommons.org/licenses/by-nc/3.0/>), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.