Annals of Global Health VOL. ■■, NO. ■■, 2017

ISSN 2214-9996/$36.00
https://doi.org/10.1016/j.aogh.2017.09.002

VIEWPOINT

Making Online Outbreak Surveillance Work for all

Yanbai Andrea Wang, JD, PhD, Michele Barry, MD, FACP
Stanford, CA

KEY WORDS epidemics, event-based internet surveillance, global health, health equity, infectious disease, surveillance, World Health Organization.

INTRODUCTION

Delayed detection and reporting in recent epidemics such as Zika and Ebola have brought renewed attention to outbreak surveillance. In particular, the use of digital technologies to crowd-source and interpret volumes of public information from the Internet for signs of outbreaks—so-called event-based Internet surveillance (EBIS)—has received great interest because of its potential for early detection. The National Academy of Medicine suggests that new information technology, which has increased surveillance capacity even in low-resource settings, be fully exploited.[1] The United Nations recommends that the World Health Organization (WHO) create an open platform to manage and analyze public data on unusual health events globally.[2] Developing the capacity for event-based surveillance is also required under the 2005 International Health Regulations (2005 IHR),[3] an international law that binds 196 states worldwide.

The growing significance of EBIS in global health raises questions concerning its efficacy, accessibility, and impact. Technical challenges, unreliable funding, institutional exclusivity, and legal shortcomings combine to generate uneven effects across the globe.

A PATCHWORK OF TOOLS

Emerging in the 1990s with the proliferation of the Internet, EBIS has since become a patchwork of overlapping tools created by a multiplicity of actors and institutions principally in the developed world. Whereas traditional indicator-based surveillance relies on structured information such as case reports and laboratory results, EBIS uses diverse sources of public information from the Internet such as news articles, messages on forums, information on health agency websites, individual investigator reports, blogs, search engine queries, and social media.

These disorganized raw signals, containing information of varied quality, are gathered in select languages. They are processed through machine translation and automated textual analysis that discards irrelevant documents and extracts key data about potentially relevant events. Human analysis may then be employed to determine which events warrant further investigation.

The most widely used EBIS tools today are Canada's Global Public Health Information Network (GPHIN), the International Society for Infectious Diseases' Program for Monitoring Emerging Diseases (ProMED-mail), the European Commission's Medical Information System (MedISys), and Harvard University's HealthMap. Each has unique strengths and weaknesses as a result of different information acquisition protocol, subject matter scope, geographic and linguistic coverage, funding base, and access rules (Table 1).

GPHIN is only accessible to organizations with a public health mandate, whereas the others are freely available to the public. GPHIN and ProMED-mail are the most human-driven systems, with expert analysts that moderate content. HealthMap has some human input, whereas MedISys is fully automated. Many more systems exist,[4] and the landscape remains volatile with substantial turnover.

All authors had access to the data and a role in writing the manuscript. The authors confirm that they have no conflicts of interest to declare.
From the Stanford Law School, Stanford, CA (YAW); and Center for Innovation in Global Health, Stanford School of Medicine, Stanford, CA (MB).
Address correspondence to Y.A.W. (yawang@law.stanford.edu).

Table 1. Four Widely Used Event-Based Internet Surveillance Tools

Tool	Owner & Funder	Access Policy	Information Acquisition Protocol	Subject Matter Scope	Geographic / Linguistic Coverage
ProMED-mail (operating since 1994)	Owner: International Society for Infectious Diseases (USA) Funder: corporate, foundation, and individual donors	Free and open to public. Reports viewable at http://promedmail.org/ or through e-mail subscription. More than 70,000 subscribers in at least 185 countries.	Began as an e-mail list for infectious disease experts to share news. More recently, has partnered with HealthMap to receive raw data feed from internet. Moderated by experts working part-time who screen, review, and investigate reports before they are posted online and distributed by e-mail to subscribers. Platform allows discussion and requests for information among subscribers.	Covers outbreaks of infectious diseases and acute exposures to toxins that affect human health.	There is 1 global network and 8 regional/linguistic networks: Portuguese (Latin America), Spanish (Latin America), Russian, French (Francophone Africa), and English (Mekong Basin of Southeast Asia, Anglophone Africa, Middle East, and South Asia).
GPHIN (operating since 1997)	Owner: Public Health Agency of Canada (Canada) Funder: Canadian government	Access restricted to public health community. Previously fee-based but has been free to organizations with public health mandate (eg, ministries of health, universities, etc) since 2009. More than 500 users: approximately 250 national users and 250 from the world community.	Automated program aggregates and assigns relevancy score to thousands of online news reports daily. Those above a threshold relevancy score are examined by a multilingual and multidisciplinary team of 11 full-time analysts, who may conduct further research for validation. An alert is sent if a potential risk is identified.	Covers all hazards: human, animal, and plant diseases; radiological and nuclear incidents; unsafe products; and natural and manmade disasters.	Monitors information in 9 languages: Arabic, English, Farsi, French, Portuguese, Russian, simplified and traditional Chinese, Spanish. Planning on adding Indonesian.
MedISys (operating since 2005)	Owner: Joint Research Centre (European Commission) Funder: European Centre for Disease Prevention and Control, Joint Research Centre	Free and open to public at http://medisys.newsbrief.eu/. Automated daily alert e-mails freely available. Newsletters are circulated once a week to subscribers (more frequently if a significant outbreak is ongoing).	Fully automated system that harvests and analyses hundreds of thousands of articles a day and also mines data from Twitter. The system performs semantic analysis to extract information such as the disease, symptoms, location, or type of public health threat implicated. Similar articles in the same language are clustered. Human input needed only for multilingual key word selection for automated textual analysis.	Covers human and animal infectious diseases; chemical, biological, radiological, and nuclear threats; and food contamination.	Monitors information in 50 languages, but functions best in common European Union languages, Chinese, Arabic, and Farsi.
HealthMap (operating since 2006)	Owner: Harvard University (USA) Funder: government, corporate, foundation, and individual donors	Free and open to public at http://www.healthmap.org/. Approximately 100,000 visits per month and 10,000 registered users. Mobile app, Outbreaks Near Me, also available to public for free download.	Automated program continuously scans internet for information on potential disease outbreaks. Machine learning system conducts text processing and classifies potential outbreaks by disease and event location. This information is then pushed to and visualized on a world map. A team of about 20 part-time student analysts review content and may correct or refine classifications.	Covers emerging infectious diseases.	Monitors information in 16 languages: English, Spanish, French, Russian, Arabic, Chinese, Korean, Japanese, Vietnamese, Bahasa (Malay, Indonesian, and some Hindi), Portuguese, German, Italian, and Thai.

GPHIN, Global Public Health Information Network; MedISys, Medical Information System; ProMED-mail, Program for Monitoring Emerging Diseases.