

Stephen L. Roberts and Stefan Elbe
**Catching the flu: syndromic surveillance,
algorithmic governmentality and global
health security**

**Article (Accepted version)
(Refereed)**

Original citation:

Roberts, Stephen L. and Elbe, Stefan (2016) *Catching the flu: syndromic surveillance, algorithmic governmentality and global health security*. *Security Dialogue*. ISSN 0967-0106

DOI: [10.1177/0967010616666443](https://doi.org/10.1177/0967010616666443)

© 2016 The Authors

This version available at: <http://eprints.lse.ac.uk/68209/>

Available in LSE Research Online: November 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Catching the Flu:

Syndromic Surveillance, Algorithmic Governmentality and Global Health Security

Abstract

How do algorithms shape the imaginary and practice of security? Does their proliferation point to a shift in the political rationality of security? If so, what is the nature and extent of that shift? This article argues that efforts to strengthen global health security are major drivers in the development and proliferation of new algorithmic security technologies. In response to a seeming epidemic of potentially lethal infectious disease outbreaks – like HIV, SARS, pandemic flu, MERS, Ebola and Zika – governments and international organizations are now using several next-generation *syndromic* surveillance systems to rapidly detect new outbreaks globally. This article analyses the origins, design and function of three such internet-based surveillance systems: 1) the Program for Monitoring Emerging Diseases, 2) the Global Public Health Intelligence Network, and 3) HealthMap. The article shows how each newly-introduced system became progressively more reliant upon algorithms to mine an ever-growing volume of indirect data sources for the earliest signs of a possible new outbreak – gradually propelling algorithms into the heart of global outbreak detection. That turn to the algorithm marks a significant shift in the underlying problem, nature, and role of knowledge in contemporary security policy.

Keywords: Algorithms, global health security, governmentality, pandemic, syndromic surveillance

Introduction

Rapid advancements in automated and digital inter-connectivity during the late 20th and early 21st centuries have spawned new ways of imagining and practicing security. In particular, a wide array of algorithmic-driven technologies now seek to harness the data-based, predictive capacities imbued in these developments in order to inform governments about the probability of threats that are as yet unforeseen. Such algorithmic technologies represent new strategic instruments of security and pre-emption – especially through their seeming capacity to infinitely amass, aggregate and transcribe unintelligible mass data sets, thus rendering visible and intelligible future-facing knowledge for the rapid identification of insecurity. The increasing integration and centrality of such automated and algorithmic technologies within a broad spectrum of contemporary security practice has already been documented – for example, in relation to the proliferation of digital algorithms for human iris detection at airports and national borders (Daugman, 2004), in relation to automated identification capacities for enhanced surveillance within human biometrics and facial recognition programmes (Introna and Wood, 2004), and – more generally – within the context of homeland securitization in the ‘War on Terror’ in the United States and elsewhere (Amoore, 2009). In fact, digital algorithmic processes appear increasingly ubiquitous within the security and surveillance strategies of the twenty-first century.

Efforts to strengthen global health security are a major, if frequently overlooked, driver for the development and international proliferation of such new algorithmic security technologies. The experience of a seeming epidemic of new epidemics over the past two decades – from HIV/AIDS, SARS and pandemic flu, through to MERS, Ebola and Zika – has begun to re-contour perspectives of insecurity, and accorded greater international political centrality to the threat that international public health emergencies can pose for populations, economies, states and infrastructures (Elbe, 2006, 2007a; Lakoff, 2008; 2015; Samimian-Darash, 2011). Already these repeated outbreak experiences have compelled many governments to make the protection of their populations and economies against the emergence of new, lethal infectious diseases an explicit part of the national security strategies (Elbe, 2009; 2010; 2014). Like so many other areas of security practice, policy discourses on global health security similarly resonate now with notions of ‘early detection’, ‘pre-emption’ and ‘response’.

Algorithms are central to these developments because they have enabled the creation of several next-generation *syndromic* surveillance systems now routinely used by governments and international organizations to rapidly detect new infectious disease outbreaks occurring around the world. In the United Kingdom, Public Health England (2015) defines syndromic surveillance as ‘the process of collecting, analysing and interpreting health-related data to provide an early warning of human or veterinary public health threats, which require public health action.’ The precise definition and application of syndromic surveillance systems still varies considerably in practice, and remains subject to scholarly debate – with some scholars even noting underlying ‘confusion about the terminology’ (Morse 2012: 9). Indeed,

Henning (2004:1-2) observes, ‘specific definitions for syndromic surveillance are lacking and the name itself remains imprecise’ – with the term presently being used to cover a wide array of early warning and outbreak detection systems. Notwithstanding those competing definitions, Morse rightly suggests that most definitions of syndromic surveillance ‘highlight the use of “non-diagnostic” data – that is, information on possible health events before, or without, definite laboratory confirmation’ (Morse 2012: 9).

Syndromic surveillance thus marks a departure from more traditional forms of public health surveillance, which tended to rely upon the reporting of *official* scientific and statistical health information to guide responses to emergent health emergencies. By contrast, digital syndromic surveillance functions through the constant, omnipresent and (near) real-time monitoring, collection and reporting of a range of *non-diagnostic* (and often open-source) data to detect *early* signals of a new infectious disease outbreak. Rather than waiting for the older, and usually lengthier, process of *direct* clinical and laboratory confirmation of a new infectious disease outbreak, syndromic surveillance systems continuously monitor a wide range of more *indirect* data – such as reports from hospital emergency departments, hospital admissions, sales of medicines from pharmacies, telephone calls to health advice providers, levels of absenteeism at school and/or workplace, etc. – for early indications that a new outbreak or even a bioterrorist attack may have occurred (Public Health Action Programme Triple S-AGE, 2010). The main idea behind the use of such ‘proxy’ data is that if a novel infectious disease breaks out, then informal signals of unusual clusters of illnesses may surface before any official clinical and laboratory analysis can be undertaken. For example, people suddenly becoming ill may start to search for unusual symptoms on internet search engines, may begin to purchase over the counter remedies, or may decide to stay home from work or school. Through closely and continuously monitoring these early, preclinical signals it may become possible to considerably speed up the process of outbreak detection – thus gaining vital time for preparing a government response.

Over the years the type (and spectrum) of data that such systems can draw upon has gradually expanded. This has given rise to a new set of syndromic surveillance systems analysing a much wider variety of non-traditional data sources for such early signs of an outbreak (Morse, 2012: 9). In particular, the past two decades have witnessed the creation of several new syndromic surveillance systems harnessing the growing availability of open source public health news and information that is now widely available and/or exchanged over the internet (Zhang, Dang, Chen, and Thurmond, 2009: 509). The wager behind these internet- and news-based syndromic surveillance systems is that new cases of unusual symptoms, illness or death occurring anywhere around the world might well become the subject of local media reports, or online discussion, long before a government can officially confirm a new outbreak and report it to the relevant international health organizations. Such syndromic surveillance systems are now

extensively relied upon by the World Health Organization for their international outbreak detection activities (Wilson and Brownstein, 2009; Zhang, Dang, Chen, and Thurmond, 2009: 509).

Yet processing the immense volume of indirect and mass data sets in near real time is a substantial challenge. That is why automated algorithmic technologies are increasingly relied upon to report and transcribe the potentially dangerous geographies of global health. Within this era of ‘Big Data’, in which continually-generating and widely accessible open-source data streams provide ‘the ability... to harness information in novel ways to produce useful insights’, the algorithm has emerged as a *credible knowledge logic* for managing the means through which information is produced, circulated, and consumed in a complex informational society (Mayer-Schönberger and Cukier, 2013:2; Gillespie, 2013:191). This article traces the origins, design and function of three novel online syndromic surveillance systems for strengthening global health security: 1) the Program for Monitoring Emerging Diseases, 2) the Global Public Health Intelligence Network, and 3) HealthMap. The articles shows each new system has become progressively more reliant upon algorithms to continuously mine an ever-growing volume of indirect data sources for the earliest signs of a possible new outbreak – propelling algorithms into the heart of global outbreak detection. That turn towards the algorithm signals a significant shift in terms of how knowledge is problematized, the kind of knowledge that is produced, and the underlying role that knowledge performs in contemporary security practices.

1. First Steps: The Program for Monitoring Emerging Diseases (ProMed-mail)

Governments wishing to protect their populations and economies against new infectious disease outbreaks want to be alerted of a new outbreak as quickly as possible. Time is considered critical for acquiring a better understanding of the nature, epidemiology and likely spread of the new disease. It is also crucial for preparing government responses and – where necessary – readying healthcare systems for a surge in patients. Yet governments have historically encountered at least two obstacles in achieving rapid international outbreak detection.

First, because new outbreaks can occur anywhere in the world, governments are effectively reliant upon the honesty (and capability) of other governments to rapidly and accurately report new infectious disease outbreaks to the rest of the international community where and when they occur. In the past many governments have been loath to be candid and forthcoming about such new outbreaks, due to fears about the possible impact of such news on trade, tourism and so forth. At the apex of the SARS epidemic emerging in southern China in 2002-2003, for example, international public health experts had to confront the spectre of Chinese state sovereignty and significant government stone-walling. Yet international health organizations such as the World Health Organization (WHO) could not officially take action until China had come forward to formally notify them of the outbreak. Such ‘political’ challenges in attaining timely and verified outbreak data in the case of SARS highlighted significant discrepancies

between the monitoring and reporting of infectious disease outbreaks, as well as an over-reliance on nationally-curated health information, within global health surveillance initiatives. The political sovereignty of other countries thus represents one pertinent obstacle to a more rapid and reliable system of global outbreak detection.

In addition to potential ‘political’ interference with rapid infectious disease outbreak reporting, a second challenge stems from the traditionally lengthy process of scientifically confirming and notifying the presence of a new infectious disease. The process of safely collecting samples of new microbes, conducting complex laboratory analyses, and feeding the information through the relevant government channels can prove a lengthy process, and also one prone to bureaucratic delays and/or errors. In reflecting upon an outbreak of poliomyelitis in Guinea in 1967/68, for example, Weir and Mykhalovisky (2010:118), argue that within these older governmental systems of health surveillance ‘[o]utbreak and epidemic control occurred in local time and place, were weakly articulated to both official and unofficial knowledge in the first weeks of onset, and were incorporated in national and global time after the fact.’ In a global context where people (and therefore also microbes) could move rapidly across the world on the backing of an international airline infrastructure, but where the reporting and surveillance mechanisms of traditional health infrastructures generally functioned within the office-hours of national disease institutes and scientific laboratories, interest emerged in developing faster detection systems that would move infectious disease outbreak surveillance closer to ‘real’ time temporalities (Lakoff, 2015).

One of the first attempts to develop such a novel outbreak detection systems was the Program for Monitoring Emerging Diseases (ProMed-mail). Tracing its origins to the fledgling days of the World Wide Web and emerging online interconnectivity in 1994, ProMed-mail was designed as an effort by the Federation of American Scientists to streamline emerging practices of global health surveillance. The central objective of this first-of-kind online surveillance system was to ‘promote communication amongst the international infectious disease community, including scientists, physicians, epidemiologists, public health professionals and others interested in infectious diseases on a global scale... [and] to participate in discussions on infectious disease concerns, to respond to requests for information, and to collaborate together in outbreak investigations and prevention efforts’ (ProMed-mail, 2010). Looking back, ProMed-mail is now regarded as the original prototype of open-source syndromic surveillance systems for global health security (Zhang, Dang, Chen, and Thurmond, 2009).

The design and purpose of ProMed-mail differed from previous systems of health surveillance since its inception. As Lawrence Madoff (2004:227) points out, a founding principle of the new online health technology was that it would remain open to *all* global sources and not just official health data produced by government health authorities. The new system would function by harnessing the power of personal computers, email and the internet to connect subscribers working in different countries. Those subscribers could then use the networked platform to exchange reports and information about new

infectious disease outbreaks that they became aware of – irrespective of where they were geographically located in the world. Critically, this networked platform meant that the ProMed-mail system could draw upon *both* official data *and* also report other types of more ‘informal’ material submitted by its subscriber base. As it evolved, the system thus gathered information from a diversity of outlets including official government health and departmental reports, but also press briefings, bulletins from international organisations, as well as professional or personal observations and other open source material of subscribed ProMed-mail members, media sources and even local rumours (Cowan, Garland, Hugh-Jones, Shimshony, Handysides, Kaye, Madoff, Pollack, and Woodall, 2006:1091).

Those incoming reports from an array of sources are subsequently analysed by expert subject area moderators within the ProMed-mail team. Those experts possess wide-ranging thematic and scientific knowledge about bacterial, viral, and parasitic diseases, as well as animal and zoonotic diseases, plant disease, epidemiology and entomology (ibid). Through the use of human analytic expertise, and the medians of the personalized computer and of the World Wide Web, the ProMed-mail system thus seeks to either verify or discredit submitted global health data and to scan the documents for accuracy and relevancy. Relevant incoming health reports are also further verified by the ‘top’ moderator on duty within the system for additional review and consideration, and with a view to assigning reports an urgency rating. An estimated 30 daily disease notification reports are received from ProMed-mail subscribers, with an estimated 20 of these reports being forwarded to the Top Moderator on duty for further verification (Madoff, 2004:229).

In the next step, moderators can then assign an urgency rating to the incoming reports. A red rating indicates a public health emergency of highest concern. In this case reports are immediately uploaded onto the system’s public website, with further notification emails being sent to ProMed-mail subscribers on the system listserv. Reports listed as yellow may or may not be uploaded to the website with discretion left to the top moderator. Items with low or no urgency are labelled green and sent on for further review and finalisation. In terms of its overarching time horizon, this system of outbreak news classification and dissemination thus enables infectious disease outbreak reporting to become a daily activity, with the ProMed-mail system publishing and uploading seven daily reports, 365 days per year, on infectious disease trends.

ProMed-mail has been incredibly successful at attracting subscribers from around the world. The new system may have only counted 40 members at its launch in 1994; but by 1997 interest in novel online open-source reporting had already witnessed the growth of the subscriber-base to 11,000 members spread across 135 countries. As the initiative approached its tenth anniversary in 2004, ProMed-mail membership had nearly tripled to more than 32,000 individuals based in 150 countries. By 2014, and 20 years following the novel launch of the ProMed-mail, the subscribership consisted of more than 60,000 members in at least 185 countries (Woodall, 1997; Madoff, 2004; ProMed-mail, 2010). The extensive

uptake of the new system is remarkable in and of itself; but it also further extends the system's geographic coverage and penetration because the system allows members to submit information that they encounter locally. The more people around the world who use the system, the more information on infectious disease outbreaks can be garnered and exchanged.

With the benefit of hindsight, ProMed-mail marked a pivotal development in the history of infectious disease surveillance as a pioneering effort to harness emergent technologies for faster and more extensive surveillance of infectious disease outbreaks across the world (Weir and Mykhalovskiy, 2010: 88). In terms of its underlying temporalities, ProMed-mail was able to move closer to the goal of attaining real-time reporting by introducing a cycle of *daily* reporting. Politically, it was far less constrained by traditional demarcations of the state and national sovereignty because it was located online and within the emergent world of the World Wide Web. Epistemically, it utilised a greater range and volume of sources by combining both official and unofficial information sources. Geographically, it achieved ever-greater coverage as its membership began to emanate from more and more countries around the world. All of those aspects combined to make ProMed-mail quite a revolutionary system of surveillance capable of creating a regime of visibility around infectious diseases not previously achievable or available within official government programmes of health surveillance. And yet ProMed-mail also remained a surveillance system still extensively reliant upon the capacity of human analysts to analyse, translate and upload all this relevant reporting of potential infectious disease outbreaks. As innovative a system as it was, ProMed-mail did *not* yet make extensive use of algorithmic technologies.

2. Harnessing the Algorithm: The Global Public Health Intelligence Network (GPHIN)

Only three years after the inception of ProMed-mail another new syndromic surveillance system called the Global Public Health Intelligence Network (GPHIN) was developed in partnership between the World Health Organization (WHO) and Health Canada. Focusing explicitly on open source media reports, the operational objective of GHPIN was 'to determine the feasibility and effectiveness of using news media sources to continuously gather information about possible disease outbreaks worldwide, and to rapidly alert international bodies of such events' (Keller, Blench, Tolentino, Freifeld, Mandl, Mawudeku, Eysenbach and Brownstein, 2009:690). GPHIN would essentially aim to constantly monitor online news and media reports from around the world (and in multiple languages) with the aim of detecting the earliest signs of a possible new infectious disease outbreak.

The idea for the GPHIN initiative can be traced back to the 1994 outbreak of pneumonic plague in India. On 23 September 1994 reports of human infections with pneumonic plague began to surface in Surat, a city located within the extremely densely populated state of Gujarat. The outbreak resulted, albeit briefly, in an unprecedented panic with serious international ramifications. The perceived rapid spread of the Surat Plague in the following days provoked an outward mass flight of an estimated quarter of the 1.5

million city-dwellers from Surat across India. The Surat plague thus seemed to capture all the fears of how a potentially lethal infectious disease could suddenly break out and then rapidly spread to other geographic areas – posing a risk to an ever-growing number of people within and beyond the borders of India.

Surat was certainly not the first outbreak of plague; but it was nevertheless striking because the outbreak was exposed, tracked and widely reported upon by an emergent, 24-hour global news media (Dutt, Akhtar and McVeigh, 2006; Pallipparambil, 2014). The emergence of increased global interconnectivity and digital media at the end of the 20th century meant that media organizations such as BBC and CNN would come to report upon the plague situation in Surat, and would widely disseminate such unofficial health knowledge to their viewers (Weir and Mykhalovskiy, 2010). Crucially, those media reports appeared to come in more quickly than many of the traditional channels of infectious disease outbreak reporting (Pallipparambil, 2014). In many respects the media effectively began to constitute an alternative, widely-accessible – albeit more informal – disease surveillance network beyond the control of national jurisdictions.

For some the experience of the Surat plague exposed the ways in which official systems of infectious disease outbreak reporting were clearly lagging behind the pace of events and media coverage. For others, however, such rapid media reporting also represented a critical opportunity to further speed up the process of international outbreak detection. If there could be an effective way of continuously monitoring such media reports in near real time, this could lead to much earlier warnings about new outbreaks than having to wait for the official diplomatic channels of disease notification to run their course. The aim of the new GPHIN system would thus be to analyse and report upon emergent and probable health data trends by processing mass collated data in the form of international media reports. The underlying data would be harnessed and filtered through novel information aggregating technologies. Unlike the earlier ProMed-mail system, the operation of GPHIN thus relied heavily on media aggregation systems – notably two models called *Factiva* and *Al Bawaba*. *Factiva* is an online multi-lingual, media content service owned by the Dow Jones & Company, whereas the *Al Bawaba* service amasses and makes available media reporting sources in Arabic and English. In employing such automated surveillance capacities, the GPHIN platform automatically scans these sources every 15 minutes for input data to be processed further by a combination of human analytics, as well as automated and algorithmic programming (Mawudeku and Blench, 2006:9).

Indeed, a crucial difference in GPHIN relates to the growing role that automated processes play in analysing the data streams. Corresponding to the rise in computing capacities and further refinement of the internet, formidable portions of the GPHIN platform became reliant upon automatic-computing and digital operation processes to supplement human infectious disease observation, classification and reporting. In addition to the design and application of web-crawling programs to continuously extract

relevant open-source health data from selected websites, GPHIN thus became the first online health syndromic surveillance model to integrate an information retrieval algorithm to further observe, classify and filter aggregated incoming data-sources. GPHIN had effectively begun to harness the power of the algorithm for speeding up the analysis of incoming reports in a way that ProMed-mail had not.

In fact, the turn towards algorithmic programming within GPHIN was the key to managing the growing volume of continually generating mass online data sets being fed into the system. Through automated computing facilities the GPHIN operating platform effectively scans and ‘pulls’ pertinent global health media reports in recurrent 15 minute intervals, 24 hours a day, seven days a week from the news aggregator systems. It then locates and retrieves from this data, specific words, expressions, phrases and syntax relevant to designed health categories within the GPHIN system – including human, animal and plant diseases, biological, chemical and radiological risks and natural disasters (Keller, Blench, Tolentino, Freifeld, Mandl, Mawudeku, Eysenbach, and Brownstein, 2009:690). It does all of this, moreover, with sources in a number of different languages. GPHIN even provides timely and routinized automated, machine translation of curated English articles into Arabic, Chinese (simplified and traditional), Farsi, French, Russian, Portuguese, and Spanish report formats.

The incorporation of such algorithmic technologies within the GPHIN platform meant that even the initial classification of how significant a new report is could now become automated. The information retrieval algorithm, based upon predefined classification by taxonomy, thus *automatically* processes and assigns an initial ‘relevancy score’ by category. This score is ‘derived from the proprietary algorithm utilizing the values attributed to the keywords and terms within the taxonomies or taxonomy it has been assigned to’ (Blech, 2008:300-302). Through this process of automated observation and calculation, incoming health data that receives an assigned ‘high’ relevancy rating are automatically loaded or published onto the GPHIN database. Articles listed as immediate concerns are forwarded to the GPHIN subscribership base (in which consumers access information based on a fee-access system), and users via e-mail while articles deemed redundant or irrelevant are ‘trashed’ and not considered for further risk analysis (*ibid*).

Overall GPHIN thus signals a move towards greater reliance on algorithmic technologies for the purposes of international outbreak detection. Its use of algorithms was critical in enabling GPHIN to achieve a speed and volume of information retrieval, processing and filtering not feasible through the use of human analytics alone. GPHIN thus became one of the first news-based syndromic surveillance systems to replace traditional human labour and analytics with algorithmic processing to construct a globalised view of infectious disease risk. The power of this new, algorithmically-based system would become evident during the SARS outbreak in late 2002 and early 2003, when GPHIN was able to identify signals within the online ‘noise’ of Chinese media sources pointing to the emergence of an unusual strain of flu an estimated 11 days after the first outbreak occurred in late 2002. Crucially, GPHIN’s detection

preceded the Chinese Ministry of Health's official notification protocol to the WHO (which only occurred on February 7, 2003), as well as the WHO's first official report on the occurrence which was released to the global public on February 25, 2003 detailing the progression of an 'atypical pneumonia outbreak' (Blench, 2008:300). GPHIN had powerfully demonstrated during the SARS outbreak how – through the careful design of new algorithms – mass sets of unofficial, open-source news data could be translated into actionable indicators of a possible emerging infectious disease threat (Securing Against Future Events: Pre-emption, protocols and publics, 2015).

3. Mapping and Visualising the Outbreaks: The Birth of HealthMap

2006 witnessed the launch of yet a third automated, internet-based syndromic surveillance system called HealthMap. Whereas both ProMed-Mail and GPHIN focused on generating and disseminating information about possible infectious disease outbreaks (albeit in very different ways), HealthMap wished to go one step further by *visualising* and *aggregating* the data in a much more easily accessible and user friendly way. HealthMap would do this through use of an online, digital and global mapping system – supplemented by a Google Maps plug-in – to neatly visualise the information flows and data-streams which could otherwise be overwhelming to the user or that could even 'obscure important elements of a disease outbreak in the era of continually-generating, mass, unintelligible online data' (Keller, Blench, Tolentino, Freifeld, Mandl, Mawudeku, Eysenbach, and Brownstein, 2009:691). The advent of these next-generation syndromic surveillance mechanisms like ProMed-mail and GPHIN (as well as other systems) meant that more and more information about emergent infectious diseases was becoming available – so much so that there was a risk that this mass of online raw data would become overwhelming and meaningless without the ability to transcribe and translate their signals accurately, reliably and accessibly (Rouvroy, 2015). HealthMap was launched in response to this new challenge of 'taming the chaos' and connecting the predictive dots within continually-generating sets of mass health data.

The HealthMap system was originally designed as a multi-stream, real-time syndromic surveillance platform that monitors and continually aggregates electronic health data on new, ongoing and emerging infectious disease outbreaks (Nelson, 2008:596). As Keller, Blench, Tolentino, Freifeld, Mandl, Mawudeku, Eysenbach, and Brownstein (2009:691) note, [t]he system integrates online outbreak data from multiple electronic and digital sources including online news wires (e.g. Google News), Really Simply Syndication (RSS) feeds, expert-curated accounts (e.g. ProMed-mail), multinational surveillance reports (e.g. from Eurosurveillance) and validated official alerts (e.g. from WHO). Like ProMed-mail and GPHIN, HealthMap thus includes an array of both unofficial sources of data as well, 'including online news aggregators, eyewitness reports, expert-curated discussions and validated official reports, to achieve a unified and comprehensive view of the current global state of infectious diseases and their effect on

human and animal health' (HealthMap, 2016). Yet in building upon the premise that such 'raw' electronic data-sources were not yet well organized or integrated (Freifeld, Mandl, Reis, and Brownstein, 2008:150), the HealthMap platform utilised automated web-crawling and web-scraping programming to scan the internet on a continual, 24 hour a day, 365 days a year basis, collecting up to 20,000 online sources per hour – all with the view of better organizing and displaying this data in a more readily accessible manner.

A further innovation of the HealthMap system thus pertains to what happens next. All of this relevant aggregated raw data is subsequently loaded onto the HealthMap operating platform, where incoming disease outbreak reports are converted into standard 'alert' formats containing the fields of 'headline', 'date', 'description' and 'info text'. This information is then forwarded into an automated classification system that strips down and simplifies data sets in order to determine the precise geographic and epidemiological information associated with each alert. Text processing algorithms further juxtapose and link incoming data alongside an integrated in-system dictionary, which contains approximately 1800 disease conditions along with 5000 global geographic patterns including countries, capital and major cities, regions etc. The incoming text data is further stripped of any non-alphanumeric content and is then automatically matched alongside the HealthMap dictionary for correlation and classification between the simplified data sets and the listed geographic locations and disease conditions automatically generated by the HealthMap system.

Following this initial classification and geocoding performed through the use of algorithms, a further automated process within HealthMap then collects and tallies the number of alerts, feeds and updates obtained online about the specific occurrence of an outbreak of infectious disease within a single geographic location. The system also uses a programmed sorting algorithm in order to determine the 'heat' or relevance based upon a 1-10 integer scale for each disease in each location. Finally, this classified, translated and ranked global health reporting is then overlaid on an interactive [online] map for user-friendly access to the original report, enabling the easy viewing of the present, real-time state of global infectious disease outbreaks occurring anywhere around the globe (Keller, Blench, Tolentino, Freifeld, Mandl, Mawudeku, Eysenbach, and Brownstein, 2009:692). Like GPHIN before it, HealthMap too thus replaces the previous surveillant procedures of human observation with the more automated, continually generating capacities of digital technology and algorithmic logics. Indeed, the core objective of HealthMap was to 'allow for greater possibilities in extracting structure algorithmically from a variety of disparate data sources' (Freifeld, Mandl, Reis, and Brownstein., 2008:151).

HealthMap's ability to do so successfully was powerfully demonstrated when it detected signs of a strange fever occurring in Macenta, Guinea on 14 March 2014. That crucial signal was detected one week before the Guinean Ministry of Health officially confirmed an outbreak of the Ebola virus that would go on to become the largest outbreak of Ebola in human history. The Ebola outbreak in West Africa had showed how HealthMap's largely automated surveillance platform, powered by innovative

algorithmic technologies, was able to capture and identify specific signals, signs or aberrations of a new outbreak within a massive data-set. It had shown how the use of carefully designed algorithms could help reconcile the double-pronged problématique of Big Data – which is both potentially telling and yet seemingly infinite and unintelligible at the same time.

More broadly, the launch of the HealthMap system represented yet another step along the continuum towards *real-time* identification of potential or emergent infectious disease outbreaks. ProMed-mail had sought to create an online, real-time forum for the exchange of open-source health information; but the system remained contingent upon the temporalities of the early and emergent World Wide Web. GPHIN had subsequently sought to expand the dimensions of near-real time temporality through the utilisation of then cutting-edge innovations in web crawling programming, retrieval algorithms and simultaneous computer-translating capacities. HealthMap, in turn, marked the first online syndromic surveillance system of the Web 2.0 era, and represented a further step in the move towards capturing the ‘*real-time*’ in health surveillance reporting. It used algorithmic technologies to further accelerate real-time temporalities of health surveillance through the online visualisation, chronological mapping and open-access presentations of geographies of risk derived from data. Functioning largely through automated computing, supported to an extent by human curation to correct misclassifications and examine geographic coverage (ibid), and fully supplemented by algorithmic programming for precise information classification and geocoding, the HealthMap project reflects the latest trajectory in the growing integration of algorithmic programming into processes for global outbreak detection.

4. Syndromic Surveillance and Algorithmic Governmentality

The evolutionary arc of these of these novel syndromic surveillance systems over the past two decades has moved algorithms to the core of global outbreak detection. As a pioneering online forum of infectious disease information exchange, ProMed-mail represented a novel response to the problem of outbreak detection utilising the emerging power of the personal computer and the internet; but it still fell short of harnessing the power of the automated algorithm. GPHIN, with its subsequent integration of informational retrieval algorithms, signalled the transition to a more specifically algorithmic tactic of surveillance by automatically processing a vast volume of news data becoming openly available through online media outlets. HealthMap, in yet a further transformation, marked the rise of whole-scale implementation of automated computing processes and algorithmic programming within a single operating platform. Such growing reliance on algorithms signals a significant shift in terms of how knowledge is problematized, the kinds of knowledge that is produced, and also the underlying role of knowledge in contemporary security practices.

4.1. The Problem of Knowledge: From Scarcity to Excess

The prominent role that algorithmic technologies now play in these syndromic surveillance systems points to a pivotal change in the underlying *problem* of knowledge. Traditional systems of health surveillance were moored within national health institutes and scientific laboratories with specialist infectious disease knowledge – frequently taking the form of population records, health and laboratory reports, and so forth. These products of health knowledge tended to be time-consuming and costly to produce, as well as being almost entirely reliant on human-observation and analytics for curation. However, they were essentially geared towards creating new knowledge for the art of governing the health of populations in a context where there was previously very little, or even no, such knowledge available.

The rapid rise of the algorithm within syndromic surveillance systems, by contrast, responds not so much to a problem of information scarcity, but – on the contrary – to an underlying problem of data excess. The core challenge for syndromic surveillance is how to process an ever increasing volume of more indirect and informal information indicating that an outbreak may have occurred. ProMed-mail became the first syndromic surveillance system to seek to address this changing problematization of knowledge through its novel use of a transnational, digital platform for information exchange. The system recognised that there had been a proliferation of multiple new data sources, including many informal ones that could be better harnessed for the purposes of rapid outbreak detection. ProMed-mail was able to successfully connect these different data sources (principally via electronic mail and the internet), but the scale of the information flow was such that it could still be managed largely through human analysis. The turn towards monitoring the much vaster information ecology of online news media in the search for possible outbreaks through the GPHIN system, however, posed an additional challenge. There were so many media reports being constantly generated that they could no longer be processed in a speedy and cost-effective way through the use of human analytics alone – necessitating the introduction of automated sorting and retrieval algorithms. The subsequent co-existence of multiple such surveillance systems, together with a proliferation of other data sources, would see a further intensification in the recourse to algorithms through the launch of HealthMap system, which was largely sustained and regulated by algorithmic processing and automated computing. The turn to algorithms in these surveillance systems thus represents an attempt to tame the ‘chaos’ of excessive and constantly generating data-flows that are impossible to process solely by utilising human analytics alone. In the move from traditional public health surveillance to syndromic surveillance, the underlying problem of knowledge has, in many ways, become exactly reversed.

4.2. Proxy Data – or ‘Knowledge Without Truth’

The extensive reliance on algorithms also generates a new *type* of knowledge for governing global health security. Much of the knowledge produced by these next-generation syndromic surveillance systems is,

strictly speaking, more of a ‘surface’ knowledge of indirect signs and proxy signals. In his lectures on governmentality, Foucault had famously discussed the rise of statistics as a new form of knowledge for governing populations. The lectures even illustrated the importance of such statistical knowledge specifically in relation to the problem of governing infectious diseases – focusing on the example of smallpox and inoculation practices of 18th century Europe (2007:10). In that case statistical science – applied to the living, biological world – sought to render visible and intelligible ‘how many people are infected with smallpox, at what age, with what effects, with what mortality rate’.

The new surveillance systems explored above, however, have used a combination of algorithms and open-source data to generate a different kind of knowledge. They are part of a knowledge regime of electronic and digitised signs, codes and signals. They produce, in the words of Antoinette Rouvroy, a kind of ‘*knowledge without truth*’. By this she means, firstly, that knowledge is no longer produced, as within previous truth regimes, through the collection, quantification and utilisation of human logic, but by ‘building upon the factual availability of enormous amounts of raw digital data’ (Rouvroy, 2012:3-7). Secondly, it means that knowledge generated by algorithmic processes increasingly escapes the previous systems of verification including the kinds of trials, tests, examinations and experiments that were viewed as essential to ‘attest to the robustness, truth, validity and legitimacy of claims and hypothesis about reality’, within previous governmental regimes of truth (ibid).

The emergence of this new kind of ‘surface’ knowledge in the area of global health security became evident in 1994 when the new ProMed-mail system began to depart from previous health surveillance systems by also processing and curating *non-official, non-clinical* open-source information. The system thus marked a considerable ‘loosening’ of the inclusion criteria for suitable outbreak data that could be considered. GPHIN subsequently utilised the emergent World Wide Web as a forum in which to communicate and disseminate an *unofficial* and largely *digitized* form of ‘surface’ knowledge gained from online media sources. The launch of HealthMap in 2006 represented yet a further step in this direction through the introduction of largely automated surveillance platform. Indeed, HealthMap deployed algorithmic processing to determine correlations between captured raw data-feed and to visualise a truly networked, digitised view of present global health trends informed entirely from patterns and signals within data-sets *contra* official statistical verification and scientific confirmations of previous systems of knowledge. Overall, those novel syndromic surveillance systems are premised less on transcribing actual cases of illness and/or outbreaks, and are more geared toward the interception of unmediated data-flows – always seeking to harness data-points and correlations towards greater intelligibility and visibility of the probable or potential. However, in the end all the indirect and non-clinical data utilised by these syndromic surveillance systems to generate early warnings may – or may not – be indicative of an actual disease outbreak.

4.3. *From the Normal to the Exceptional*

The rise of the algorithm in these syndromic surveillance systems also signals a shift in the underlying *role* that knowledge performs in contemporary security practices. In presenting the epistemic transformation between a disciplinary economy of power towards that of a governmental economy of power in his *Security, Territory, Population* lecture series, Foucault sought to make distinct the particular ways in which populations are ‘normalized’ through the identification and modification of risk groups (Foucault, 2007:63). Unlike a disciplinary economy of power whereby ‘an ideal norm is imposed from the outside, and the distinction between the normal and the abnormal is undertaken according to this ‘artificial’ norm (ibid), the concept of normalization within Foucault’s governmentality once more was very heavily influenced by the dynamics of statistical analysis and of the cultivation of an average to determine abnormality and relative risk. Accordingly, normalization – as described within Foucault’s understanding of governmentality – constituted ‘the plotting of the normal and the abnormal, of different curves or normality, and the operation of normalization consists in establishing an interplay between these different distributions of normality’ (2007:63). Fearnley (2008) has shown that such processes of normalization were central within previous governmental strategies of public health, especially through the establishment of regulatory mechanisms aimed at the development of an average or an equilibrium within the overarching governance and optimisation of population health (2008:1617). However, those past systems of public health and the governance of risk were also frequently limited by costs, external conditions, and most significantly limitations within existent possibilities of knowledge (ibid).

The syndromic surveillance systems analysed here depart from the centrality of the norm and from the optimal statistical distribution for governing the health of populations. Rather than seeking to plot the normal versus the abnormal to determine the norm through interplay of facts, statistics and figures, algorithmically driven syndromic surveillance systems constantly survey in order to detect, determine and report upon that which constitutes the *exception* within the governance of infectious diseases. This divergence from traditional public health surveillance was already evident in the rationale for designing the Program for Monitoring Emerging Diseases (ProMed-mail), which – from its inception – sought to locate and report upon *emergent* infectious disease outbreaks, as opposed to determining averages and relative norms of existing disease trends within populations. GPHIN too was orientated not towards measuring and assessing the global burden of infectious disease amongst populations, but specifically to capturing the emergence of *new* infectious disease *outbreaks*. The same holds true for the algorithmic processes that capture, match and upload online mass-data sets to HealthMap. HealthMap too seeks to visualise aberrations not averages, as illustrated by the syndromic surveillance system’s initial detection of a ‘mysterious fever’ occurring in Macenta, Guinea on 14 March 2014.

Collectively these syndromic surveillance systems are thus beginning to re-contour understandings of the norm and the average through what Rouvroy (2015), highlights as the non-

selectivity and exhaustive nature of ‘Big Data’ and algorithmic processing. Where the cultivation of the norm within governmentality involved the plotting of statistical points and seeking to bring what was considered abnormal into the fold of the ‘norm’, the idea of the average within algorithmic governmentality disappears completely with the capacity for automated computing and algorithmic logic to effectively capture and present all relevant points of data, no matter how distant. Unlike the previous generation of statistical knowledge, whereby points or data which deviated too far from a central or common finding were disregarded, the infinite scope and capacity of algorithmic programming means that even the most isolated or singular of points can now be taken into greater analysis, specification and account (ibid). In the end, the rise of the algorithm within the context of these new syndromic surveillance systems signals not only a change in the nature and problem of knowledge within contemporary security practices, but also in its underlying role and purpose.

Conclusion

The three syndromic surveillance systems analysed in this article differ considerably. Yet viewed collectively they also show that algorithmic programming has become progressively more central to the practice of global health security. As the earliest of the three outbreak detection systems, ProMed-mail was the only system *not* to explicitly utilise algorithmic programming. Nonetheless, it represented an early attempt to harness and curate the knowledge available from freely available and exchangeable, open-source health data – thus marking a significant deviation from the more traditional utilisation of statistical practices and official health knowledge in health surveillance. ProMed-mail may have still relied on human analytics rather than algorithmic programming, but it laid the operational groundwork from which a more advanced integration of algorithmic technologies would soon become possible.

Such integration of algorithmic programming into syndromic surveillance was subsequently achieved through the development of the Global Public Health Intelligence Network (GPHIN). As one of the first digitally-interconnected, semi-automated health surveillance system, GPHIN sought to integrate leading machine translation and automated/digital processes to develop an online health system with a ‘real-time’, 24 hours a day, 7 days a week surveillance capacity (Mawudeku and Blench, 2006:7-11; Weir and Mykhalovskiy, 2010: 84-85). It explicitly introduced an information sorting and retrieval algorithm in response to the growing volume and complexity of information available on the internet during the early days of GPHIN. Here the launch of the GPHIN initiative also coincided historically with the turn of the millennium – a time during which personal computer usage, consumption of online information, and expanding access to the World Wide Web increased considerably.

HealthMap signalled yet a further step towards a fully automated surveillance platform. Sustained and informed even more extensively by algorithmic programming, HealthMap requires only minimal human analytics for precise geocoding, infectious disease classification, outbreak alert report curation,

and interactive visualisation of ongoing infectious disease trends globally. Indeed, and of the three systems to emerge between 1994 and 2006, HealthMap represents the most extensive and sustained effort yet to harness the power of the algorithm for the purposes of strengthening global health security. The iterative evolution of these new syndromic surveillance systems –from ProMed-mail, via GPHIN, through to HealthMap – has thus pushed algorithmic technologies to the centre of international outbreak detection.

The turn to algorithms for strengthening global health security signals a considerable shift in the underlying problem, nature, and role of knowledge within contemporary security practices. The initial introduction of the algorithm responds to an underlying problem of knowledge no longer characterised by a situation of a scarcity of data relating to probable outbreaks of infectious disease; but rather by an excess of constantly expanding and generating data mass that may contain patterns, associations and correlations indicative of yet-unforeseen public health emergencies. Indeed, the very development of these new syndromic surveillance systems is predicated upon the conviction that existent forms of knowledge generation within health security are insufficient for addressing contemporary concerns regarding the threat of globalised pandemic and infectious disease outbreaks (Lakoff, 2015). Through the introduction of algorithmic technologies, moreover, syndromic surveillance systems then generate a new and quite different kind of knowledge – more of a ‘surface’ knowledge of indirect digital signs and signals that may enable faster responses and earlier interventions. Nor is the function of this new knowledge even still that of normalization as described in Foucault’s influential account of governmentality. Whereas normalization sought to establish a particular norm through ‘the plotting of the normal and the abnormal, of different curves or normality, and the operation of normalization consists in establishing an interplay between these different distributions of normality’ (2007:63), syndromic surveillance systems move towards the detection of the *exceptional* or the aberrant within mass data sets.

Do all of these changes also point to a deeper transformation in the political rationality of security? On the one hand, all three systems considered here still remain continuous with the central aspects of governmentality described by Michel Foucault. Foucault famously referred to governmentality as ‘the ensemble formed by institutions, procedures, analyses and reflections, calculations, and tactics that allow the exercise of this very specific, albeit very complex, power that has the population as its target, political economy as its major form of knowledge, and apparatuses of security as its essential technical instrument’ (2007:108; Collier and Lakoff, 2015:26). The surveillance systems analysed in this article remain continuous with that notion of governmentality in that they too are all geared towards securing the welfare of populations – in this case against the death and illnesses that could be caused by new infectious diseases (Elbe, 2009; Elbe, Roemer-Mahler, and Long, 2014). It would thus be premature to speak of a complete transformation, or even break, in the underlying rationality of security.

At the same time, we have seen that the close integration of algorithmic technologies into processes of global outbreak detection does signal a distinct shift or inflection within Foucault’s

influential understanding of governmentality – especially in terms of how security knowledge is today problematized, understood, and the wider role it performs. Antoinette Rouvroy has usefully coined the term ‘algorithmic governmentality’ to capture the emergence of a new governmental rationality made possible by the growing complexity of information technology – one that ‘utilises algorithmic processes of data-aggregation, translation and prediction’ and that is ‘chiefly oriented towards the neutralisation of potentiality (Rouvroy, 2009). The syndromic surveillance systems analysed here constitute evidence that such a specifically algorithmic governmentality is also beginning to manifest itself more fully in the area of global health security. More than that, they show that attempts to strengthen global health security – with its dual concern about the threat posed by pandemics and bioterrorist attacks – have actually been a key driver behind the international development, enhancement and proliferation of new algorithmic security systems.

As with debates about the utility of ‘Big Data’ more generally, pertinent questions still remain about how effective, accurate and reliable these new syndromic surveillance systems are overall. This issue emerges particularly clearly in relation to *Google Flu Trends*. The system sought to improve early detection through the monitoring of ‘health-seeking behaviour in the form of queries to online search engines, which are submitted by millions of users around the world each day’ (Ginsberg, Mohebbi, Matthew, Patel, Brammer, Smolinski, and Brilliant, 2009:1012)’. The idea behind the system was that as people began to suffer from the flu, they would be likely to search the internet for information on flu symptoms and remedies. It might thus be possible to trace the patterns of such search activities, and to aggregate this data into ‘historical logs of online web search queries’, in order to predict the onset and scale of the flu outbreak (ibid). Initially, the system attracted a lot of positive attention and acclaim after reports that it was considerably quicker than the Centres for Disease Control and Prevention (CDC) in identifying the onset and scale of seasonal flu the United States (Ginsberg, Mohebbi, Matthew, Patel, Brammer, Smolinski, and Brilliant, 2009). However, it later emerged that the system was not nearly as powerful or accurate when in 2012, when the system ‘identified a sudden surge of flu cases but overstated the amount’ (Mayer-Schönberger and Cukier, 2013:2-3). More recently, there has thus been prominent coverage of the ‘failure’ of the digital surveillance system *Google Flu Trends* to accurately predicting oncoming seasonal influenza arcs (Butler, 2013; Lazer, Kennedy, King, and Vespignani 2014). The uneven experience with *Google Flu Trends* thus serves as a pertinent reminder that the accuracy and effectiveness of these new surveillance systems remains an open question.

In the end, however, it is not just the effectiveness of such systems that matters. Beyond their accuracy, the very existence and introduction of such new systems can also have wider political ramifications. In the case of global health security there is evidence that the introduction of new syndromic surveillance systems is already changing the way states are identifying outbreaks and exercising their sovereignty. The revised *International Health Regulations* (2005), for example, officially

authorized the utilisation of such *unofficial* health data generated by these surveillance models to further inform and guide responses to emergent and probable public health emergencies. Indeed, as early as 2001 Health Canada and the WHO entered into an agreement that GPHIN would supply WHO with monitoring data while the WHO would engage in verification processes with its member-state contacts, officially consolidating the GPHIN surveillance system within the operational frameworks of the WHO's *Global Outbreak Alert Response Network* (GOARN) (Heymann and Rodier, 2001;2004). In that sense, syndromic surveillance systems have formed a significant part of the global outbreak surveillance landscape for more than a decade now.

The existence of these new systems will likely also bear on the deliberations of governments regarding their decisions about whether and how quickly to report new infectious outbreaks to the rest of the international community. David Fidler (2004:144), argues that such 'unofficial health information and global alerts now act as a global health governance pincer that squeezes the state's sovereign decision of whether to report outbreak information and to cooperate with the WHO and other countries.' Governments considering whether or not to rapidly declare a new outbreak to the international community will now need to weigh up the chances that others will become aware of the outbreak through such syndromic surveillance systems, or that they may even be aware of an outbreak already. Although it is difficult to determine the precise motivations and actions of governments, evidence is emerging that the introduction of such syndromic surveillance systems appears to be having an effect. For example, a quantitative assessment in the timeliness of infectious disease outbreak reporting (Chan, Brewer, Madoff, Pollack, Sonricker, Keller, Freifeld, Blench, Mawudeku, and, Brownstein, 2010) points to a shortening of the interval between detection and public communication of outbreaks since the introduction of these new systems. In either case, however, the progressive integration of algorithms into global outbreak detection is already giving rise to a more nuanced and diversified global health surveillance regime in which, as Weir and Mykhalovskiy (2010:104) argue, '[s]overeign states did not disappear at the turn of the 21st century, but they are now reactive to the presence of a transnationalized system of outbreak communication that they themselves have politically and legally authorized.'

References:

- Amoore L (2009) Algorithmic War: Everyday Geographies of the War on Terror. *Antipode* 41(1): 49–69.
- Blench M (2008) *Global Public Health Intelligence Network (GPHIN)*. Available at: <http://www.mt-archive.info/AMTA-2008-Blench.pdf> (accessed 01 July 2014).
- Butler D (2013) When Google got flu wrong. *Nature* 494(7436):155–156.

- Chan E, Brewer T, Pollack M, Sonricker A, Keller M, Freifeld C, Blench M, Mawudeku A, and Brownstein J (2010) Global capacity for emerging infectious disease detection. *Proceedings of the National Academy of Sciences* 107(50):21701–21706.
- Collier S and Lakoff A (2015) Vital Systems Security: Reflexive Biopolitics and the Government of Emergency. *Theory, Culture & Society* 32(2):19–51.
- Cowan P, Garland T, Hugh-Jones ME, Shimshony A, Handysides S, Kaye D, Madoff LC, Pollack MP, and Woodall J (2006) Evaluation of ProMed-mail as an electronic early warning system for emerging animal diseases: 1996-2004. *Journal of American Veterinary Association* 229(7):1090-1099.
- Daugman J (2004) How Iris Recognition Works. *IEEE Transactions on Circuits and Systems for Video Technology* 14(1): 21–30.
- Dutt A, Akhtar R and McVeigh M (2006) Surat plague of 1994 Re-examined. *Southeast Asian Journal of Tropical Medicine and Public Health* 37(4):755-760.
- Elbe S (2006) Should HIV/AIDS be securitized? The ethical dilemmas of linking HIV/AIDS and security. *International Studies Quarterly* 50(1): 119-144.
- Elbe S (2007) Our epidemiological footprint: The circulation of avian flu, SARS, and HIV/AIDS in the world economy. *Review of International Political Economy* 15(1):116–130.
- Elbe S (2009) *Virus Alert: Security, Governmentality, and the AIDS Pandemic*. New York: Columbia University Press.
- Elbe S (2010) Hagglng over viruses: the downside risks of securitizing infectious disease. *Health Policy and Planning* 25(6): 476-485.
- Elbe S (2014) The Pharmaceuticalization of Security: Molecular Biomedicine, Antiviral Stockpiles and Global Health Security. *Review of International Studies* 40(5):919-938.
- Elbe S, Roemer-Mahler A and Long C (2014) Securing circulation pharmaceutically: Antiviral stockpiling and pandemic preparedness in the European Union. *Security Dialogue* 45(5):440–457.
- Fearnley L (2008) Signals come and go: Syndromic surveillance and styles of Biosecurity. *Environment and Planning* 40(7):1615–1632.
- Fidler D P (2004) *SARS, governance and the globalization of disease*. Basingstoke: Palgrave Macmillan.
- Foucault M (2007) *Security, Territory, Population, Lectures at the Collège De France*. Edited by Michel Senellart. New York: Picador.
- Freifeld C, Mandl K, Reis B, and Brownstein J (2008) HealthMap: Global Infectious Disease Monitoring through Automated Classification and Visualization of Internet Media Reports. *Journal of the American Medical Informatics Association* 15(2):150–157.
- Gillespie T (2013) The Relevance of Algorithms. In Gillespie T, Boczkowski P, and Foot K (eds.) *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: MIT Press: 167-194.
- Ginsberg J, Mohebbi M, Patel R, Brammer L, Smolinski M, and Brilliant L (2009) Detecting influenza epidemics using search engine query data. *Nature* 457(7232):1012–1014.

HealthMap (2016) Available at: <http://www.healthmap.org/site/about> (accessed 01 January 2016).

Henning KJ (2004) What is Syndromic surveillance? *Centers for Disease Control: MMWR* 53:5–11.

Heymann DL and Rodier GR (2001) Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases. *The Lancet Infectious Diseases* 1(5):345–353.

Heymann DL and Rodier GR (2004) Global surveillance, national surveillance, and SARS. *Emerging Infectious Diseases* 10(2):173-175.

Introna L and Wood D (2004) Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems. *Surveillance & Society* 2(2/3):177-198.

Keller M, Blench M, Tolentino H, Freifeld C, Mandl K, Mawudeku A, Eysenbach G, and Brownstein J (2009) Use of Unstructured Event-Based Reports for Global Infectious Disease Surveillance. *Emerging Infectious Diseases* 15(5):689–695.

Lakoff A (2008) The Generic Biothreat, or, How We Became Unprepared. *Cultural Anthropology* 23(3):399–428.

Lakoff A (2015) Real-time biopolitics: the actuary and the sentinel in global public health. *Economy and Society* 44(1):40–59.

Lazer D, Kennedy R, King G, and Vespigani A (2014) The parable of Google flu: Traps in big data analysis. *Science* 343(6176):1203–1205.

Madoff LC (2004) ProMED-mail: An Early Warning System for Emerging Diseases. *Clinical Infectious Diseases* 39(2):227–232.

Mawudeku A and Blench M (2006) *Global public health intelligence network (GPHIN)*. Available at: <http://www.mt-archive.info/MTS-2005-Mawudeku.pdf> (accessed 03 October 2015).

Mayer-Schönberger V and Cukier K (2013) *Big Data A Revolution That Will Transform How We Live, Work and Think*. London: John Murray Publishers.

Morse S (2012) Public Health Surveillance and Infectious Disease Detection. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science* 10(1):6-16.

Nelson R (2008) HealthMap: the future of infectious diseases surveillance? *The Lancet Infectious Diseases* 8(10):596.

Pallipparambil G (2014) The Surat Plague and Its Aftermath. *Insects, Disease, and History*: 1–4.

ProMed-mail (2010) Available at: www.promedmail.org (accessed 10 September 2014).

Public Health Action Programme Triple S-AGE (2010) Syndromic Surveillance Systems in Europe. Available at: <http://www.syndromicsurveillance.eu/systems-in-europe> (accessed 02 January 2016).

Public Health England (2015) Syndromic surveillance: Systems and analyses. Available at: <https://www.gov.uk/government/collections/syndromic-surveillance-systems-and-analyses> (accessed 14 December 2015).

Rouvroy A (2009) Technology, Virtuality and Utopia: Governmentality in an Age of Autonomic Computing. In: Hildebrandt M and Rouvroy A (eds.) *Automatic Computing and the Transformations of Human Agency: Philosophers of Law meet Philosophers of Technology*. Routledge: 1–32.

Rouvroy A (2012) The end(s) of critique: data-behaviourism vs. due process. In: Hildebrandt M and De Vries E (eds.) *Privacy, Due Process, and the Computational Turn. Philosophers of Law Meet Philosophers of Technology*. Routledge: 1–19.

Rouvroy A (2015) Algorithmic governmentality: a passion for the real and the exhaustion of the virtual, Transmediale- All Watched Over by Algorithms, Berlin. Available at: <https://www.youtube.com/watch?v=4nPJbC1cPTE> (accessed 04 April 2015).

Samimian-Darash L (2011) Governing through time: preparing for future threats to health and security. *Sociology of Health & Illness* 33(6):930–945.

Securing against Future Events: Pre-emption, protocols and publics (2015) Available at: www.securityfutures.org (accessed 23 June 2015).

Weir L and Mykhalovskiy E (2010) *Global Public Health Vigilance: Creating a World on Alert*. 1st edn. New York: Routledge.

Wilson K and Brownstein J (2009) Early detection of disease outbreaks using the Internet. *Canadian Medical Journal Association* 180(8): 829-831.

Woodall J (1997) Outbreak meets the Internet: Global epidemic monitoring by ProMED-mail, ProMed-mail Publications. Available at: <http://www.promedmail.org/aboutus/publications> (accessed 06 October 2013).

World Health Organization (2005) International Health Regulations, Second Edition. Available at: http://www.who.int/publications/2008/9789241580410_eng.pdf?q=international (accessed 23 June 2015).

Zhang Y, Dang Y, Chen H, and Thurmond A (2009) Automatic online news monitoring and classification for syndromic surveillance. *Decision Support Systems* 47: 508-517.