Big Data, Algorithmic Governmentality and the Regulation of Pandemic Risk

Stephen L ROBERTS*

This article investigates the rise of algorithmic disease surveillance systems as novel technologies of risk analysis utilised to regulate pandemic outbreaks in an era of big data. Critically, the article demonstrates how intensified efforts towards harnessing big data and the application of algorithmic processing techniques to enhance the real-time surveillance and regulation infectious disease outbreaks significantly transform practices of global infectious disease surveillance; observed through the advent of novel risk rationalities which underpin the deployment of intensifying algorithmic practices to increasingly colonise and patrol emergent topographies of data in order to identify and govern the emergence of exceptional pathogenic risks. Conceptually, this article asserts further howthe rise of these novel risk regulating technologies within a context of big data transforms the government and forecasting of epidemics and pandemics: illustrated by the rise of emergent algorithmic governmentalties of risk within contemporary contexts of big data, disease surveillance and the regulation of pandemic.

I. Introduction

The threat of pandemic illness represents one of the most serious and complex security challenges in an era of cascading global risks.[1] Over the course of the past decade, the governing of health and population security have been increasingly problematised by a seeming epidemic of epidemics,[2] of emergent and re-emergent pathogens. The 2014 West Africa Ebola Virus epidemic killed thousands and cost billions of dollars to eventually contain. The outbreak of Ebola was quickly followed by globalised epidemics of the Zika virus and the Middle East respiratory syndrome (MERS), the resurgence of Ebola in the Democratic Republic of the Congo and, in February 2018, a novel, highly

* Email: s.l.roberts1@lse.ac.uk.

[1] As noted by Rasmussen (2001), appropriations of *risk* in the field of international relations have tended to conflate the concept of risk with those of danger and threat and have failed to spell out the conceptual difference between security and risk. Within the context of this article, risk is conceptualised as the estimation of a dangerous future. It implies a specific relation to the future, a relation that requires a monitoring of the future, an attempt to calculate what the future can offer and to develop novel solutions to manage contingency. While traditional security practices were essentially based on the possibility of empirically identifying and assessing threats, risk reshapes the relation to what is knowable by introducing uncertainty and the unknowable at the heart of security governing processes, thus deploying *technologies of risk* to identify, calculate and mitigate "incalculable" threats such as terrorism and pandemic. See further C Aradau et al, "Security, Technologies of Risk, and the Political" (2008) 39(2–3) Security Dialogue 147.

[2] J Bartlett, "An Epidemic of Epidemics", *Medscape Infectious Diseases*, 28 February 2014.

European Journal of Risk Regulation, 10 (2019), pp. 94–115
doi:10.1017/err.2019.6

infectious strand of avian influenza, H7N4 emerged in China, prompting fears of a new global pandemic.[3]

Correspondingly, within the global informational society of today, exponential growths in the volume, scope and scale of digital "big data" sources now afford unprecedented new opportunities for the collection, processing and translation of digital data reservoirs to understand, to analyse, and *regulate* forthcoming risks. In an era in which 10 billion messages per day are exchanged on Facebook, where 3.5 million text messages are sent every minute, and where 2.5 quintillion bytes of data have been produced every day since 2016,[4] the generation of new streams of infinite big data sources now represents an epistemic break from previous strategies of risk management which were characterised by an *incompleteness* of information.[5] Increasingly, the application of big data analytics and the proliferation of algorithmic processing techniques is evident across a broad spectrum of security practices for the monitoring and regulation of a multitude of emerging risks, ranging from natural disasters and climate change, to terrorism, civil wars, and the insecurity of global financial markets.

In an era of heightened global pandemic vigilance, whereby complex infectious disease outbreaks constitute critical risks to networked global economies and populations, new rationalities to regulate yet unforeseen public health emergencies have emerged with increasing intensity over the past two decades within global health security frameworks. In representing a critical departure from previous systems of surveillance and risk forecasting these emergent techniques of risk assessment seek to apprehend data reservoirs, and to *secure with algorithms*,[6] in order to obtain accelerated insights of probable pandemic outbreaks via processing of "big data" to make sense of the torrents of available information.[7]

How do these intensified recourses to the "promises" of "big data" transform practices of disease surveillance and the real-time regulation of pandemic risk in the 21st century? Does the increased centrality of *big data* and *algorithms* within disease surveillance systems signify the emergence of distinct novel rationalities within the politics of disease surveillance and risk regulation? If so, how do these shifts towards harnessing big data transform attempts to regulate the emergence of pandemic risks in an era of circulating pathogens and proliferating digital data sources?

The aim of this article is to demonstrate how, within an era of data overload and data potential, global disease surveillance practices have moved towards an ever intensified recourse to the integration of big data and algorithmic processing as increasingly vital surveillance functions, to accrue and *make sense* of infinitely expanding digital mass-data sets to forecast the earliest indications of potential pandemic risks. In tracing the rise and proliferation of these ideologies and practices of "big data" over the past two decades

[3] World Health Organization, "Human Infection with Avian Influenza Virus", < www.who.int/csr/don/22-february-2018-ah7n4-china/en/ > (accessed 13 February 2019).

[4] IBM < www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/ > (accessed 13 February 2019).

[5] Big Data & Risk Conference, Concordia University < www.bigdatariskconference.wordpress.com > (accessed 13 February 2019).

[6] L Amoore and R Raley, "Securing with Algorithms: Knowledge, decision, sovereignty" (2017) 48(1) Security Dialogue 3.

[7] B Rieder, "Beyond Surveillance: How Do Markets and Algorithms 'Think'?" (2017) 31(8) Le foucaldien 1.

within frameworks for disease surveillance and pandemic regulation, this article develops and presents two key contributions to contemporary analyses of big data and the regulation of risk.[8] First, the article argues critically, that intensified efforts towards harnessing big data and the application of algorithmic processing techniques to enhance the real-time surveillance of infectious disease outbreaks produce new and emergent transformations in the practice of infectious disease surveillance, which increasingly colonise and patrol the realm of the *digital* in order to understand, anticipate and predict forthcoming risks of the *pathogenic* realm. Second, this article further asserts that the rise of these novel rationalities of disease surveillance within a context of big data transforms the government and forecasting of epidemics and pandemics: illustrated by the rise of emergent algorithmic governmentalities of risk, within contemporary contexts of big data, disease surveillance and the regulation.

II. Data, surveillance and the forecasting of pathogenic risk

1. Disease outbreaks, surveillance and the "avalanche" of statistical numbers

From the catastrophic Black Death, which killed 75% of the European population in the 1300s, to the 1918 "Spanish flu" global pandemic, which killed 5% of the world's population, to the emergence of HIV/AIDS, Ebola and SARS, the control and regulation of disease outbreaks, epidemics and pandemics have been vital to the survival and continuity of communities and population groups for centuries. Information-gathering practices of disease surveillance, which involve the collection, analysis and interpretation of large volumes of health-relateddata, have underpinned efforts to regulate and control the spread of pathogens and infection across borders and populations and have further informed and guided nascent disease control measures from the application of quarantines, segregations, vaccinations and public hygiene campaigns.

The origins of modern infectious disease surveillance programs date back several centuries in Europe and North America to the late 18th century, in which the collection and logging of health and population data first emerged as a central feature of government.[9] Critically, new data sources afforded by the avalanche of statistical numbers[10] of the modern era, made possible for the first time the surveillance of whole populations through manual and routine collection and classification of numerical and demographic data,[11] which could be applied to determine disease control responses to regulate the emergence of epidemics.

[8] The ideology of "big data" is predicated upon the belief that provided one has access to massive amounts of raw data, one might be able to anticipate most phenomena of the physical and digital worlds thanks to relatively simple algorithms. See further A Rouvroy, "The end(s) of critique: data-behaviourism vs. due process" in M Hildebrandt and K de Vries (eds.), *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology* (Routledge 2013).

[9] M Foucault, *Security, Territory, Population Lectures at the Collège de France 1977–1978* (Picador 2007); I Lowrie, "Algorithmic rationality: Epistemology and efficiency in the data sciences" (2017) Big Data & Society 1.

[10] I Hacking, "Biopower and the avalanche of printed numbers" (1982) 5(3–4) Humanities in Society 279.

[11] D Armstrong, "The rise of surveillance medicine" (1995) 17(3) Sociology of Health & Illness 393.

With the further refinement and advancement of national health infrastructures, the accrual and processing of statistical health data by clinicians, statisticians and epidemiologists became an essential function of the modern state in the battle against emerging epidemics and pandemics. By 1901, all US States had laws requiring reporting of selected communicable diseases to local authorities,[12] and by 1913 the US Public Health Service recommended that State and Territorial health authorities should submit weekly statistical summaries reporting infectious diseases.

The founding of the World Health Organization (WHO), from 1948 onward, accelerated new imperatives towards the transparent and timely sharing of disease surveillance data from its member states in the interest of controlling and addressing increasingly complex and emergent infectious disease outbreaks which threatened to upend global circulations of capital and populations, including outbreaks of cholera, plague, poliomyelitis and smallpox which occurred globally with growing intensity and scope from the mid-20th century onward.

2. Data scarcity, emergent digital data sources and health security

Despite early advancements to regulate and control the emergence of pandemic illness throughout the 20th century, expanding infectious disease surveillance frameworks have been historically and operationally problematised by a marked *scarcity* of health surveillance data and an *incompleteness* of epidemic intelligence required to estimate and respond to the coming patterns of infectious disease outbreaks across an increasingly networked yet vulnerable global system.

Problematically, established surveillance practices of data accrual, analyses and dissemination of epidemic intelligence were notorious for severe time-lags, often focused only on very specific disease profiles, and offered little opportunity for non-state actors to play any role in official epidemic intelligence gathering,[13] although in many incidences, the press and media did operate as social actors capable of disseminating news of outbreaks, independently of the WHO and its member states.[14] Furthermore, the accessing and dissemination of critical infectious disease data throughout the early life course of theWHO was frequently stonewalled or obfuscated by the sovereign concerns of its member states over the negative impacts stemming from reporting an occurring outbreak.

This inability of WHO authorities to access requisite epidemic intelligence further exacerbated deadly and destructive outcomes exemplified by virulent outbreaks of poliomyelitis and cholera which occurred in Guinea in 1967 and 1970.[15] Further still, facing a continued crisis of surveillance data for the rapid reporting of emerging epidemic and pandemic risks, established health surveillance practices fell steadily out of pace with the globalisation of (re)emergent diseases from the 1970s onward.

Occurring, however, amid these increasingly globalised public health emergencies, waves of "computerisation" from the mid-20th century onwards, as observed by

[12] SB Thacker and RL Berkelman, "Public Health Surveillance in the United States" (1988) 10 Epidemiological Review 164.

[13] S Davies and J Youde, "The IHR (2005), Disease Surveillance and the Individual in Health Politics" (2013) 17(1) International Journal of Human Rights 11.

[14] L Weir and E Mykhalovskiy, *Global Public Health Vigilance: Creating a World on Alert* (Routledge 2010) 75.

[15] ibid.

Rieder,[16] have involved considerable changes and extensions in terms of what computers *can do*, and the growing power of the computer has subsequently inserted computing technologies and capacities into almost all domains of life. Timeliness, ease and access to critical surveillance data and epidemic intelligence increased substantially for example due to computing innovations, with the advent of electronic public health and hospital record systems in the 1970/80s.[17] The increasing commonplace of computer processing and digital storage of health data launched the first computerised disease surveillance system, the *Réseau Sentinelles* (Sentinelles Network) in France in 1984,[18] marking new overtures to enhanced surveillance and disease-tracking capacities offered by expanding digitisation and computation processes.

Growing trends towards the incorporation of digitisation and computation into infectious disease surveillance intensified further during the 1990s, whereby heightened anxieties surrounding the potentialities of bioterrorism and intensifying global pandemic prospects merged with the unprecedented generation of open-source digital data following the launch of the internet. This proliferation of digitised and unofficial epidemic intelligence disseminated across the World Wide Web gave rise to a novel form of infectious disease surveillance in the early digital age: *syndromic surveillance*.

Syndromic surveillance may be defined 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".[19] It thus represented a critical break away from the traditional clinical surveillance processes of past eras, grounded in manual data accrual and the generation of statistical forecasting, and built alternatively upon intercepting and processing the rapidly expanding data sources produced from the rise of the Internet to generate new techniques for the regulation of rare and emerging disease threats.[20] The advent of syndromic surveillance represented, therefore, the first recourse to apprehending expanding big data streams for the digital tracking of disease outbreaks in an era increasingly preoccupied with emergent pandemic risks. Early systems, including ProMed-mail, launched in 1994, and the Global Public Health Intelligence Network (GPHIN) which launched in 1997, demonstrated great initial operative successes in harnessing then-nascent big data sets to forecast emergent public health emergencies. Amongst these successes were ProMed-mail's reporting of an outbreak of Ebola in Kikwit, Zaïre in 1995,[21] and GPHIN's widely-cited early identification of the SARS coronavirus in Southeastern China in late 2002.[22]

[16] Rieder, supra, note 7, 1.

[17] L Simonsen et al, "Infectious Disease Surveillance in the Big Data Era: Towards Faster and Locally Relevant Systems" (2016) 214(4) The Journal of Infectious Diseases 380.

[18] A Valleron, "A Computer Network for the Surveillance of Communicable Diseases; The French Experiment" (1996) 11(76) American Journal of Public Health 1289.

[19] Public Health England < www.gov.uk/government/collections/syndromic-surveillance-systems-and-analyses > (accessed 13 February 2019).

[20] M Flear, "'Technologies of Reflexivity': Generating Biopolitics and Institutional Risk To Supplement Global Public Health Security" (2017) 8(4) EJRR 7.

[21] Weir and Mykhalovskiy, supra, note 14, 155.

[22] M Dion et al, "Big Data and the Global Public Health Intelligence Network" (2015) 41(9) Canada Communicable Diseases Report 209.

Following the early reporting successes of these novel surveillance systems, practices of digitised syndromic surveillance were rapidly and enthusiastically integrated into the early response and outbreak notification infrastructures of the WHO including its technical surveillance network, the Global Outbreak Alert and Response Network. Between 1997 and 2001, an estimated two-thirds of infectious disease surveillance data and epidemic intelligence received by WHO authorities came first from these digitised surveillance platforms, rather than from traditional surveillance reporting provided by WHO member states and ministries of health.[23] Critically, the emergence of these syndromic surveillance systems which sought to enhance the surveillance and reporting of pandemic risk in the late 20th century also facilitated a novel wider turn in practices of surveillance towards the development of new digital surveillance technologies, and the implementation of accelerated data-processing capacities to address contingent pandemic risks.

III. Big data and the regulation of pandemic risk

Collectively, while such syndromic surveillance systems served as early forerunners of risk technologies which sought to harness big data, the first two decades of the 21st century would witness a further incorporation of big data parsing and algorithmic processing as core operating components in digital disease surveillance systems, to reduce, as Tyler Reigeluth writes, the relative indeterminacy of the [uncertain] future to a predictable and computational sequence of *that which is to come*.[24]

Illustrating these contemporary leaps towards accelerated disease surveillance fed by big data and informed by algorithmic data-processing, in 2009, a novel and highly infectious strain of the Influenza A virus (H1N1) emerged in a rural village in Veracruz, Mexico. Weeks before the emergence of H1N1, internet giant Google published an article in the science journal *Nature*,[25] in which it successfully predicted the coming arc of seasonal winter influenza several weeks in advance of the Centers for Disease Control and Prevention, through the automated matching and comparing of 50 million of the most common search queries and terms associated with influenza from 2003/08. Critically, Google was able to track cases and locations of suspected cases of influenza in real-time, producing accelerated epidemic intelligence detailing an expanding global public health emergency. This novel approach to parsing mass-data sets via automation and algorithmic processing to inform the digital surveillance of emergent epidemics proved to be a more useful and timely indicator than government statistics, with their natural reporting lags.[26]

[23] D Heymann and G Rodier, "Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases" (2001) 1(5) The Lancet 345.

[24] T Reigeluth, "Why data is not enough: digital traces as control of self and self-control" (2014) 12(2) Surveillance & Society 245.

[25] J Ginsberg et al, "Detecting influenza epidemics using search engine query data" (2009) 457 Nature 1012.

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

1. Enter the algorithm: from data *incompleteness* to data *excess*

The famous parable of H1N1 and Google Flu Trends signified therefore new transformative shifts towards the problematisation of *data incompleteness*, which has historically existed at the core of infectious disease surveillance processes to regulate contingent pandemic threats in the epochs which preceded the rise of "big data". It is estimated, for example, that humanity accumulated 180 EB[27] of data between the invention of human writing systems and the year 2006. Between 2006 and 2011, however, owing to the advent of the Web 2.0 era, the total amount of existing data sources proliferated tenfold and reached 1,600 EB,[28] and yet problematically, 80% of all digital data infinitely generated across the planet exists in an *unstructured* form and cannot be presented or comprehended in tables of relational databases.[29]

Thus, within the rise of big data, a new problematisation of data emerges for the practice of global disease surveillance, not one of *data incompleteness*, which hallmarked previous informational eras, but one of *data excess*. As such, the tale of Google and H1N1 signified the emergence of novel risk assessment practices for pandemic vigilance in a world increasingly submerged by big data: (1) to capture as much data as possible about *everything*, to use in novel ways – for assessing infectious disease distributions, to tracking business trends, mapping crime patterns, and analysing web traffic; and (2) to recognise further that humans simply cannot do that kind of risk analysis unassisted.[30] Within this expanding era of digital information complexity in which aggregated data sources can be both *meaningful* and *meaningless*, the digital algorithm emerges across big datascapes as a strategic and "objective" purveyor of crucial epidemic intelligence.

The years which followed the H1N1 big data experiment and the piloting of Google Flu Trends have witnessed a proliferation of novel disease surveillance technologies and new practices orientated towards apprehending torrents that can confer a competitive advantage in the forecasting and tracking of local or international outbreaks.[31] In addition to Google Flu Trends, Twitter, the online news and social networking service which launched in 2006, and which produces over 6,000 tweets per second, has become a site extensively surveyed by the *algorithmic gaze* of digital disease surveillance, and intensified efforts have been employed to determine how the big data bulk of billions of tweets curated by the social media platform and its user-base can be aggregated, leveraged and analysed via algorithmic data-processing to produce real-time surveillance data of occurring or probable disease outbreaks.

Since 2008, numerous health surveillance initiatives have employed evolving algorithmic logics to process and filter billions of tweets from Twitter users to produce

[27] EB = Exabyte.

[28] L Floridi, "Big Data and Their Epistemological Challenge" (2012) 25 Philosophy & Technology 435.

[29] IBM, supra, note 4.

[30] M Andrejevic and K Gates, "Big Data Surveillance: Introduction" (2014) 12(2) Surveillance & Society 185.

[31] C Garattini et al, "Big Data Analytics, Infectious Disease and Associated Ethical Impacts" (2017) Philosophy & Technology 1.

real-time detections of infectious disease outbreaks including human and avian influenza patterns,[32] the Ebola Virus Disease,[33] and through the detection and monitoring of HIV cases and infections using Twitter network data.[34] While the big data reservoirs of the Twitter platform have been recognised as a new strategic commodity for public health surveillance and the monitoring of infectious diseases, relatively little attention has been focused, as underscored in a study by Conway,[35] on the development of ethically appropriate regulatory approaches to working with these new data sources, whereby a single researcher can automatically process hundreds of millions of public tweets.[36]

Further still, in demonstrating increasingly advanced algorithmic capacities to apprehend and process big data to monitor and report upon the emergence of potential pandemics, on 14 March 2014, a text-processing algorithm designed by the digital disease surveillance system HealthMap identified an unusual media report from the big data bulk of hundreds of thousands of websites and online media outlets which detailed the emergence of strange haemorrhagic fever (*étrange fièvre*), observed in Macenta Prefecture, Guinea.[37] The posting of this outbreak alert on the HealthMap website occurred over a week in advance of the official confirmation by the Guinean Ministry of Health on 22 March 2014, which verified that an outbreak of Ebola was occurring in several areas across the country. This early identification of an aberrant health event identified by a digital algorithm sparked further attention in how increasingly automated and digitised modes of disease surveillance, informed by big data, could be leveraged to improve the early detection and response to public health emergencies of international concern.[38] In the wake of the West Africa Ebola epidemic, an article in the *Lancet Infectious Disease* asserted further that encouragement of countries to adopt, integrate and apply these evolving digital surveillance technologies to track and regulate the movement of pandemic risks should be a global public health priority.[39]

While the piloting of new disease surveillance technologies and techniques by Google, Twitter and HealthMap represent distinct enterprises, the common thread which underpins these three examples is the extent to which health surveillance practices now seek to apprehend and process big data via algorithms to illuminate infinite datascapes to extract meaning and to derive certainty of forthcoming risks from them.[40] Within the

[32] K Kim, "Use of Hangeul twitter to track and predict human influenza infections" (2013) 8(7) PloS ONE; J Li and C Cardie, "Early Stage Influenza Detection from Twitter" (2014) arXiv preprint.

[33] M Odlum and S Yoon, "What can we learn about the Ebola outbreak from tweets?" (2015) 43(6) American Journal of Infection Control 563.

[34] SD Young, C Rivers and B Lewis, "Methods using real-time social media technologies for detection and remote monitoring of HIV outcomes" (2014) 63 Preventative Medicine 112.

[35] M Conway, "Ethical Issues in Using Twitter for Public Health Surveillance and Research: Developing a Taxonomy of Ethical Concepts from Research Literature" (2014) 16(12) Journal of Medical Internet Research.

[36] ibid.

[37] Africaguinee.com, < www.healthmap.org/ebola/#timeline > (accessed 13 February 2019).

[38] A Anema et al, "Digital surveillance for enhanced detection and response to outbreaks" (2014) 14(11) The Lancet Infectious Disease 1035.

[39] ibid.

[40] M Pasquinelli, "Anomaly Detection: The Mathmatization of the Abnormal in the Metadata Society" (2015) Paper presented at Transmediale Berlin, 1–10.

post-Ebola landscape of disease surveillance and pandemic monitoring, the capacity of algorithms to automatically detect potential health emergencies based on digital data accrual and processing indicates further that data-driven analytics as new practices of knowledge creation for addressing uncertain events are on the rise – not only in economic contexts, but in the fight against terrorism and transnational crime, border controls, large-scale events,[41] and the tracking and regulation of pandemic risk in the 21st century.

IV. Governmentality and the securing of future pandemics

Amid the widespread deployment of big data processing technologies for tracing the next global pandemic in an era hallmarked by the Edward Snowden revelations, heightened concerns for data-privacy and protection, and the Facebook/Cambridge Analytica data-harvesting scandal, how does the harnessing of big data and the unleashing of the algorithm transform risk rationalities embedded at the core of disease surveillance practices? How do these new methods of big data accrual and the production of digitised surveillance alter the way in which previous systems cultivated, produced and disseminated knowledge for the governing of infection and disease? Does the rise of big data-suffused forms of disease surveillance suggest the emergence of novel techniques for the assessment and regulation of pandemic risk in the era of the digital?

Scholarship within the field of surveillance studies traditionally has traced much influence from the conceptual frameworks presented by Michel Foucault, most widely from his illustrations of the relationship between technologies of security and population, at the moment of the emergence of the issue of population in the era of governmentality.[42] Significantly, the writings and lectures of Foucault in the late 1970s have offered conceptual tools for understanding the significance of this problematisation of disease for the contemporary government of life.[43] Specifically, Foucault's 1978 lecture-series *Security, Territory, Population* in which he discussed the governing of populations in tandem with the emergence of naturally occurring phenomena, including pandemic illness, through processes of surveillance, normalisation and calculative security practices, are well established in discourses of public health.[44] Across the social sciences, Foucault's theorisations of governmentality and the proliferation of public health security apparatuses have influenced productive

[41] M Leese, "The new profiling: algorithms, black boxes, and the failure of anti-discriminatory safeguards in the European Union" (2014) 45(5) Security Dialogue 494.

[42] Foucault, supra, note 9. In the *Security Territory, Population* lecture-series Foucault provides the definition of governmentality as: the ensemble formed by the institutions, procedures, analyses and reflections, the calculations and tactics that allow the exercise of this very specific albeit complex form of power, which has as its target population, as its principal form of knowledge political economy, and as its essential technical means apparatuses of security.

[43] A Lakoff, "Real-time biopolitics: the actuary and the sentinel in health" (2015) 44(1) Economy and Society 40.

[44] D Couch et al, "Public health surveillance and the media: a dyad of panoptic and synoptic control" (2015) 3(1) Health Psychology and Behavioural Medicine 128.

scholarship within the fields of health, security and surveillance highlighted by the works of Martin French and Gavin Smith,[45] Stephen Collier and Andrew Lakoff,[46] Marilou Gagnon and Adrian Guta,[47] Jeremy Youde,[48] Stuart Elden,[49] and Mark Flear,[50] amongst others.

Post-Foucault, the advent of the digital era, the subsequent deluge of data produced by the new digital innovations, and growing recourses to automated computing technologies in surveillance processes, however, usher forward new theorisations and sites of investigation in an attempt to understand contemporary social and technological transformations in surveillance and the regulation of risk. This article thus far has demonstrated how over the past two decades infectious disease surveillance systems, long problematised by a historical *incompleteness* of data information, have moved to increasingly integrate practices of big data accrual and algorithmic processing into core functions of disease-tracking logics to regulate the emergence of global pandemic risks. These novel risk assessment and regulation techniques which now oscillate towards the harnessing and processing of big data reservoirs as a key object of government have moved "big data" and algorithmic processing techniques correspondingly to the core of new surveillance practices to detect the next pandemic before it occurs, thus rendering uncertain futures knowable, and thus *governable.*

And yet, while a number of critical security scholars have drawn attention to how novel practices of risk estimation and forecasting are employed in the digital age and with what political effects,[51] including Claudia Aradau and Tobias Blanke,[52] Louise Amoore,[53] and Rita Raley,[54] equally broad transformations occurring at the interfaces of infectious disease surveillance and the regulation of pathogenic risks remain critically overlooked. In light of these transformative shifts in seeking to digitally assess and regulate circulating pandemic risks, this article employs and further extends the conceptual work of Rouvroy's *algorithmic governmentality,* as a contemporary analytical framework in which to highlight and comprehend the proliferation of such data-driven disease surveillance devices in an era of pandemic uncertainty.

---

[45] M French and G Smith, "'Health' surveillance: new modes of monitoring bodies, populations, and polities" (2013) 23(4) Critical Public Health 383.

[46] S Collier and A Lakoff, "Vital Systems Security: Reflexive Biopolitics and the Government of Emergency" (2015) 32(2) Theory, Culture & Society 19.

[47] M Gagnon and A Guta, "Mapping HIV community viral load: Space, power and the government of bodies" (2012) 22(4) Critical Public Health 471.

[48] J Youde, *Biopolitical Surveillance and Public Health in International Relations* (Palgrave MacMillan 2010).

[49] S Elden, "Governmentality, calcuation, territory" (2007) 25 Environment and Planning D: Society and Space 562.

[50] Flear, supra, note 20.

[51] C Aradau and T Blanke, "The (Big) Data-security assemblage: Knowledge and critique" (2015) 2(2) Security Dialogue.

[52] ibid.

[53] L Amoore, "Data derivatives: On the emergence of a security risk calculus for our times" (2011) 28(6) Theory, Culture & Society 24.

[54] Amoore and Raley, supra, note 6.

1. [Algorithmic] governmentality

The rise of modes of *algorithmic governmentality* are situated within the epoch of "big data", and it is indeed the contemporary fixation towards harnessing and processing big data to pre-empt the emergence of risk which informs and intensifies new methods of algorithmic government in the digital era. As Antoinette Rouvroy writes, "[o]perations of collection, processing and structuration of [big] data for purposes of data-mining and profiling, helping individuals and organizations to cope with circumstances of *uncertainty*…have become crucial to public and private sector's activities in domains as various as crime prevention, marketing, entertainment, and health management".[55] Rouvroy's conceptualisation of algorithmic governmentality represents a significant extension of Foucault's original premise of governmentality in emphasising how the era of "big data" provides unprecedented opportunities for data aggregation, analysis and correlation, and in doing so marks a distinct transition away from traditional statistical perspectives and governing via processes of normalisation which were central to the construction of knowledge systems for addressing the contingent.[56]

In contrast then, to the finely tuned, collected and curated "statistics of the state", which were formulated and produced by the statistician and the clinician within laboratories and national health institutes, emergent forms of algorithmic governmentality, powered by big data, seek alternatively to apprehend latent, open-source unstructured data, to render correlations visible and to discover patterns in data-warehouses.[57] In seeking now to regulate Foucault's notion of the *milieu*, through the mining of digital data reservoirs and the piloting of algorithmic technique, the impacts of the computational turn on practices of governmentality, as suggested by Rouvroy, are far from trivial.[58] Within a contemporary era fixated with the regulation of future security risks, submerged in proliferating data sources and accelerated by digital and technological innovations, *algorithmic governmentality* indicates the ascendancy of:

> "an apolitical rationality founded on the automated collection, aggregation, and analyses of Big Data to model, anticipate, and pre-emptively affect possible behaviours….it is an unprecedented mode of government fuelled mostly with signals (raw data and meta-data)… and emancipated from the yoke of representation, from all kinds of tests, even statistical norms, in data-mining and machine learning, even the quantitative notion of the average – and the normative figure of the average man – disappear".[59]

As Rouvroy's illustrative analysis here suggests, recent convergences and infusions of technological innovations (namely, the proliferation of big data) with risk assessment and security practices have produced new governing and regulation techniques which markedly diverge from previous processes and strategies of assessing, understanding and mitigating forthcoming security risks. Hallmarks of this new regime of algorithmic government thus include: the quantification and harnessing of "big data" via algorithmic

[55] Rouvroy, supra, note 8, 1.

[56] A Rouvroy and T Berns, "Gouvernementalité algorithmique et perspectives d'émancipation" (2013) 177(1) Réseaux 163.

[57] Rouvroy, supra, note 8, 5.

[58] ibid.

[59] Rouvroy and Berns, supra, note 56; A Rouvroy, "Algorithmic Governmentality: a passion for the real and the exhaustion of the virtual" (2015) Transmediale – All Watched Over by Algorithms 1–2.

processing (in contrast to the statistically-guided processes of risk assessment of preceding epochs); a distinctive divergence towards collecting and processing big data to detect that which constitutes the aberrant or the exceptional, derived from heterogeneous data streams (rather than previous strategies of governing from that of the norm or the average); and which is informed and guided by distinct methods of data collection and knowledge production which emphasise the continuous, real-time and cost-effective collation of digital data sources and automated production of risk assessments via new recourses to artificial intelligence including algorithmic processing. Significantly, while the ascendancy of modes of algorithmic governmentality in the digital era have been previously extended and applied to the regulation of risk within legal/jurisdictive realms, and in reference to the digitisation of market and economy, the rise of big data suffused infectious disease surveillance technologies for the tracking of pandemic over the past two decades signifies further another critical arena which has been steadily colonised by the logics of an algorithmic governmentality in seeking to regulate the emergence of pandemic risks via novel overtures to apprehending and processing big data. Correspondingly, the ascendency of this new governmentality of risk regulation appears to produce distinct and divergent transformations in the practice of health security and surveillance in a number of critical aspects.

a. The "problem" of data: from scarcity to excess

The increasingly salient role of big data and algorithmic programming within the practice of infectious disease surveillance over the past two decades indicates a new and shifting rationality towards the problematisation of *knowledge* within security apparatuses oriented towards the identification of future risks. As discussed, traditional practices of infectious disease surveillance which centralised the manual collection and accrual of health-related data have long served as sentinels in the identification of forthcoming epidemic and pandemic risks across populations. Problematically however, traditional surveillance processes of data analysis and the production of population health reports and logs were highly time-consuming, costly to produce, were entirely reliant on human observation and analytics for curation and dissemination, and further were frequently plagued by informational gaps, misclassifications and processing delays,[60] most critically during public health emergencies.

And yet, in seeking to regulate the emergence of global pandemic risks in the 21st century through the processing and assessment of big data, the data-driven practices of digital disease surveillance systems analysed in this article diverge significantly from the central governmental tenants of previous apparatuses of security described by Foucault. The problematisation which these data-driven technologies of risk seek to reconcile is not how to construct intelligence of emergent risks and challenges from scarce data-sources but contrastingly, how to navigate *oceans of data* to exploit and detect that which constitutes the abnormal or the contingent from infinitely generating torrents of big data.

This new relation to risk stands in sharp contrast, of course, to the government and security of epidemics including smallpox and the mechanism of variolation (inoculation

[60] M Klompas and D Yokoe, "Automated Surveillance of Health Care-Associated Infections" (2009) 48 Clinical Infectious Diseases 1268.

with the virus of smallpox)[61] which Foucault famously described within *Security, Territory, Population*. The parable demonstrates how, at the threshold of the governmental state, practices of statistical estimation and forecasting, culled from the collection and analysis of new data forms, gave rise to new practices of health security: inoculation campaigns which sought to regulate and manage the arcs and distributions of epidemic outbreaks within populations. The example illustrates further how, within emergent practices of governmentality, statistical forecasting was utilised to develop knowledge and security practices for the government of life-species, where previous such knowledge was lacking or did not exist.

Subsequently, the "deluge of big data" rather than the "avalanche of statistical numbers" poses new problematisations to the accessing and production of data in the digital era. Instead of seeking to collect and amass scarce statistical information to inform and develop a future calculus of probabilities, the belief in the growth of "big data", as highlighted by Antoinette Rouvroy[62] "is that, provided one has access to massive amounts of raw data (and that the world is actually submersed by astronomical amounts of digital data) one might be able to anticipate most phenomena of the physical and digital worlds thanks to relatively simple algorithms". The problem of data and of knowledge generation within new algorithmic modes of government therefore constitutes a quandary of **excess**, and it is at the core of this new problematisation of knowledge of big data in which the algorithm and automated data processing techniques emerge as new instruments of government in forecasting and regulating pandemic risk in the digital age.

The advent and rise of new digital surveillance technologies for the tracking of disease outbreaks from the 1990s onward has demonstrated this steady recourse to algorithmic modes of risk regulation with growing intensity over the past two decades. For example, in 1997 responding to the increasingly complex informational ecology of the World Wide Web, GPHIN became the first digitised disease surveillance system to turn towards new logics of the algorithm, implementing new techniques of digital algorithmic management in order to search, aggregate, and cross-reference large data sets[63] to produce new forms of epidemic intelligence from proliferating *unstructured* and *unintelligible* big data sets. What is more, the now infamous history of Google Flu Trends and its forecasting of the 2009 H1N1 influenza pandemic through the automated parsing of 50 million of the most common search queries and terms associated with influenza from 2003/08 demonstrates further the reverse relationship and problematisation of knowledge in the digital era of disease surveillance and pandemic vigilance. The initial widely-cited success of Google Flu Trends in tracking the spread of the global movement of H1N1 saw the intensification and use of other digital disease tracking platforms including Biocaster and EpiSPIDER, both which utilise algorithmic programming to mine and process big data, and from 2010 onward, a flurry of surveillance research began to orientate towards how algorithmic processing of Twitter's infinite reservoir of big data tweets and its content could be mined and processed for enhanced influenza detection and surveillance.

[61] Elden, supra, note 49, 566.

[62] Rouvroy, supra, note 8, 1.

[63] D Boyd and K Crawford, "Critical Questions for Big Data" (2012) 15(5) Information, Communication & Society 663.

Algorithmic governmentality thus exhibits a new rationale of risk regulation and analysis in the digital era. Rather than the sole generation of "statistical" knowledge for the governing of emergent natural phenomena, where previous knowledge was scare, the growing utilisation of the algorithm as a processor of big data, seen for example within these ascendant disease surveillance systems seeks rather to address and "tame the chaos" of infinite sources of potential security knowledge found within digital data, and to produce actionable intelligence derived from patterns of big data to inform and guide responses to regulate potential pandemic risks. In a current world saturated by data as well as circulating pathogens, and imbued with the transformative potentialities of "big data", the digital algorithm emerges within larger health security and risk assessment paradigms as an essential, yet practical new instrument of security and forecasting, enabled, as Rocco Bellanova has underscored, with the capacities to automatically capture and process data from multiple sources, using calculations that humans and socio-political institutions are by and large no longer able to understand or master.[64]

b. Shifting topographies of surveillance: from the clinic to the data-warehouse

Further still, the rise of the informational paradigm of "big data", and the emergence of algorithmic modes of regulation, significantly transform geographies and spaces in which practices of risk estimation and surveillance are imagined and enacted. Returning once more to *Security, Territory, Population*, Foucault for example went to great lengths to illustrate the emergence of natural phenomena as new problems of government technique; Foucault's three illustrations included the problem of urban organisation (the town), the problem of food scarcity (grain), and the problem of recurring epidemics (contagion).[65]

Accordingly, it was in response to the emergence of these uniquely modern contingencies that governmental perspectives shifted towards the piloting of mechanisms or dispositifs of security which sought to regulate the emergence of natural phenomenon through calculation, determination of probability and the *delimitation of phenomena within acceptable limits.*[66] Through Foucault's analysis the physical setting of the town emerges as a new space in which practices of security and surveillance are enacted to ensure the positive *circulation* of capital, economy and goods. Within these new security practices of the governmental era, population, with its datum of variables, subsequently emerges as an object of knowledge for the governing of the contingent, as well as a unit of analysis of the physiocrats and economists.[67]

It is here, at the birth of the governmental state and the deployment of new population security practices, in which the spaces of modern bureaucracies came to assume a vital security/surveillance function. In 1890, French sociologist Gabriel Tarde[68] stated how with new generalised practices of statistical forecasting and calculation "[t]he public journals will become socially what our sense organs are vitally. Every printing office will

[64] R Bellanova, "Digital, politics and algorithms: Governing digital data through the lens of data protection" (2017) 20 (3) European Journal of Social Theory 330.

[65] Foucault, supra, note 9, 62–71.

[66] ibid.

[67] ibid.

[68] G Tarde, *The Laws of Imitation* (Holt Publishing 1903) cited in Pasquinelli, supra, note 40, 1.

become a mere central station for different bureaus of statistics. At present, statistics is a kind of embryonic eye, like that of the lower animals which sees just enough to recognise the approach of foe or prey". The bureaus of statistics, therefore, within the governmental state, with its generation of population data detailing varied accounts of employment, unemployment, criminality and insanity, health and infections served centrally as the primary producers of population specific data which could be leveraged to inform security practices which governed the emergence of the contingent including epidemic and pandemic illnesses, produced in clinics, laboratories, and national health ministries.

Contrastingly, the deluge of big data advance and produce new critical sites in which data to guide and inform practices to regulate the contingent flourish. Unlike within previous governmental eras where nothing of the material world remained untouched by the statistician,[69] the infinite generation of open-source data over the past two decades has produced new ecologies in which modes of algorithmic government seek to patrol, apprehend and colonise in the governing of the emergence of risk: data-warehouses. Illustrating this radical shift in governmental topographies, in response to growing streams of infinite digital data, Google established its first data-warehouse in 1998, known as Google Cage,[70] and this digital infrastructure sought to store, accumulate, catalogue and process torrents of big data for innumerable ends including market and consumer analysis, the monitoring and forecasting of social tendencies and environmental anomalies.[71] The launch of Google Cage would be followed by a proliferation of new data storage and processing infrastructures within a digital informational era of increasingly data complexity. With the rise of big data, these gigantic datacentres represented, therefore, new monopolies of collective data, and thus emergent sites of risk calculation, surveillance and identification via processes of data-mining and the unleashing of algorithmic processing logic.

Seen here then, the proliferation of the data-warehouse in an era of algorithmic governmentality contrasts critically with the rise of the statistical bureau of preceding eras. Unlike the bureaus of statistics, which generated knowledge of the contingent from data manually collected and processed from the *physical* and *pathogenic* realms, the production of knowledge within the data-warehouse consists of disparate and heterogeneous data, processed and presented by increasingly sophisticated algorithms, pulled from the realm of the *digital* to produce actionable knowledge on risks emergent from the *pathogenic* and *physical* realms. Moreover, unlike the curation of statistics within the governmental bureaus of health institutes, laboratories and clinics which were subjected to systems of verification, hypothesis and testing, the production of digital epidemic intelligence discovered in the data-warehouse, representing critical epistemic shifts in the rationalities which underlie knowledge production practices for pandemic risk analysis, have an aura of "pure" knowledge – not on the basis of traditional criteria of authenticity, historical coherence or critical apperception, but merely on the merits of immediate operationality, plasticity, flexible adaptation … and immediate availability.[72]

[69] Hacking, supra, note 10.

[70] Pasquinelli, supra, note 40, 2.

[71] ibid.

[72] Rouvroy, supra, note 8, 5.

What is more, the rise of algorithmic modes of governance and risk analysis, which seek to leverage and process big data for pre-emptive security, produces not only new infrastructures and topographies of data in which to be surveyed, seen as the data-warehouse and datascapes, but further accelerates the ascent and centralisation of new health security actors in an era of data deluge: data monopolists. Exhibited prominently by the advances and experimentation on the part of Google in its intensifying stake in digital disease surveillance, transnational data corporations including Google and Twitter now constitute increasingly authoritative health security actors in the production and dissemination of digital epidemic intelligence.

The parable of Google Cage for instance, illustrates not only how these increasingly powerful data corporations have innovatively produced new digital ecosystems for expedited and enhanced risk analysis and assessment, but moreover in era marked by austerity and pervasive threats, the data monopolists emerge as the select few within modes of algorithmic government who possess the advanced technical/logistical, and indeed financial, means requisite to collate, process, understand and produce new forms of digitised risk analysis and epidemic indicators in the surveillance of the next public health emergency.

Understood collectively then, the rise and promises of "big data" have increasingly afforded the availability of digitally-curated risk alerts and assessments extracted from heterogeneous data streams, expanding across topographies of data and stored increasingly across proliferating data-warehouses. Indeed, this type of data, and methods of data extraction via algorithm differ markedly from the zenith of the statistical epoch of government, hallmarked by the statistical bureaus of health institutes and scientific laboratories. By contrast, within the rise of new algorithmic rationalities of governance and risk analysis, data-warehouses and associated datascapes have emerged as critical new infrastructures in which to employ powerful algorithmic processing for the accelerated production of risk assessment, fed by infinite sources of potentially indicative data, which are pulled from datascapes, produced within data-warehouses, and are afforded by a range of increasingly salient new contemporary health security actors including computer scientists, private corporations and data monopolists.

c. Data accrual: from normalisation to "collect it all"

The rise of big data analysis and the ascendancy of the algorithm further transform existent rationalities in how and what sources of data are selected, accumulated and processed towards estimating uncertain horizons. Within previous practices of disease surveillance, the manual accrual, processing and dissemination of disease and population-specific data gave rise to new understandings of patterns of infections, epidemics and pandemics, and subsequently enabled the inauguration of new regulatory measures. As recent as 1980, the WHO utilised such precise generation and application of statistical population-focused health data in collaboration with its member states in the development of mass inoculation campaigns and surveillance operations in health efforts to immunise and eventually eradicate the human smallpox virus through its comprehensive Smallpox Eradication Programme.

In presenting the emergence of new anticipatory and calculative methods which accompanied the transition to the governmental state, Foucault illustrated the centrality of processes of *normalisation* as practices of security which sought to regulate and monitor populations in the modern era. Processes of normalisation, conceptualised within governmentality by Foucault ,consisted of "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".[73]

Illustrating these security logics, Foucault utilised examples of how statistical calculation and processes of normalisation were employed to inform the *medicine of epidemics*,[74] in the determining which population and age groups were to be inoculated against the smallpox virus when analysed in relation to a coefficient of probable morbidity or probably mortality.[75] Public health and the practice of traditional health surveillance of epidemics and pandemics were thus normalising styles of governance and security, built around systems of information collection of data about life, death, and disease – but also limited by costs, external conditions and the limited possibilities of knowledge.[76]

Contrastingly, the rise of new techniques which exploit big data re-contours governmental perspectives and the centrality of establishing the average or the norm within digital practices of infectious disease surveillance. The rise of new unprecedented modes of algorithmic regulation in the digital era, inspired, as Rouvroy writes, "by new fears of imminent catastrophes and of new regulative principles such as precaution, risk minimisation, detection and anticipative evaluation",[77] exhibits marked new logics in how data sources are accessed, leveraged and processed in the contemporary forecasting of emergent security risks. Above all, instruments of algorithmic government which seek to regulate the emergence of pandemic risks in the digital era are technologies of quantification.

Unlike previous practices of security in the governmental state, which sought to employ statistical calculation to determine an average or understanding of a normal or acceptable distribution of a phenomena within a given population – and to govern from this understanding, algorithmic governmentality and the rise of big data processing systems are devoid of any relation to the average or the norm.[78] Rather than trying to determine an average or understanding of a normal or acceptable distribution of a disease or outbreak within a population, the aim of these new data processing systems is only-ever maximal data accrual, no longer to exclude anything that does not fit the average but [rather] to avoid the unpredictable.[79]

The rise of new surveillance systems in an era of big data therefore depart from the centrality of the norm and of the optimal arc of distribution, which Foucault had

[73] Foucault, supra, note 9, 63.

[74] ibid.

[75] ibid.

[76] L Fearnley, "Signals Come and Go: Syndromic Surveillance and Styles of Biosecurity" (2008) 40(7) Environment and Planning A 1167.

[77] Rouvroy, supra, note 8, 10.

[78] Rouvroy and Berns, supra, note 56, 4.

[79] ibid.

previously illustrated as a new technology for the governance of population security and the contingent. Unlike in previous governmental eras where statistical data was selectively accumulated and processed, and whereby points or patterns which deviated too far from the central or common finding were disregarded, the dramatic recent shift towards a pervasive "collect it all" security mentality, now means that even the most isolated or singular of aberrations or exceptional points can be taken into greater analysis and account in the assessment of future-situated risks.

Within these new big data-oriented mentalities of government, previous methods and understandings of normalisation which governed and regulated the emergence of future contingencies are progressively supplemented with emergent intensified capacities to apprehend and harness the volume, variety and velocity of "big data" to process data and produce information about the incalculable. Algorithmic governmentality directly looks to detect that which constitutes the *exceptional* or the abnormal through the correlation and observation of diffuse data collected and processed in a variety of heterogeneous contexts.[80]

Emergent techniques of risk analysis and regulation which employ algorithms to "collect it all" and to produce alerts and warning about incalculable exceptional events can be further located within digital disease surveillance systems of the 21st century, highlighted by the launch of the digital surveillance system GPHIN. Distinctly, the function and objective of the GPHIN system was not the generalised assessment and reporting of population health, but rather the digitised prototype sought to derive and generate knowledge of potential pandemics from complex and diversifying digital data-streams. It became the first digitised health surveillance platform to implement the sorting and foreshadowing capacities of a retrieval algorithm to observe and detect anomalies by mining and scanning a myriad of digital data-sets from a variety of sources.[81] As the first kind of health surveillance technology to employ this new data collection and processing logic via algorithm, succeeding digital platforms, including HealthMap, Google Flu Trends, BioCaster, EpiSPIDER and Twitter, have further sought to enhance the monitoring of pandemic threats and the enhancement of disease surveillance capacities through the employment of big data processing via algorithms to produce epidemic intelligence on exceptional public health episodes, including H1N1, Ebola, MERS and the Zika virus.

Collectively then, the proliferation of big data-driven disease surveillance systems represents a critical new departure from the government of the norm, originally described by Foucault. What remains vital here to emphasise with illustration and situating of the rise of these algorithmic rationalities and techniques of disease surveillance within big data contexts is not to state or seemingly imply that security perspectives which seek to govern at the interface of the norm or average can now be dismissed or rendered obsolete or passé. Contrastingly, what can be observed with the rise of novel modes of algorithmic governance are new technologies/rationalities of risk analysis exclusively oriented towards the identification of exceptions, outliers and aberrations, through new modes of security analysis and observation including pattern recognition in big data sets,

[80] Rouvroy, supra, note 8, 6.

[81] S Krasmann, "Imagining Foucault. On the Digital Subject and 'Visual Citizenship'" (2017) 23 Foucault Studies 16.

performed by algorithms. In recalling Foucault, it can be observed that the rise of these new rationalities of risk, sustained by the algorithmic identification of outliers, does not expel or render redundant, previous mechanisms of security which were based on normalisation, but rather emerge and increasingly work in tandem with existing infrastructures within expanding and intensifying assemblages of government, security and risk regulation.

As Claudia Aradau and Tobias Blanke[82] have argued within their recent compelling accounts of anomaly detection and algorithmic subjects of security, new risk assessment practices which accentuate anomaly detection or the exceptional security risk speak directly to the promise of "big data" to compute data at scale and find patterns and correlations that could reveal "unknown unknowns" or the "needle" in the digital haystack.[83] It is then the promise of "big data" which facilitates the rise of an expedited and cost-effective logic of security and risk governance with extended digital capacities to seek, to assess and to render visible and knowable the coming of exceptional security challenges, divorced from previous references to the norm. In the context of intensified practices of disease surveillance, this new logic, predicated on a pervasive "collect it all" approach to data quantification, can only be rendered feasible and operational with the continuous analysis and parsing of generating digital data sets via algorithms.

Within emergent modes of algorithmic governance, oriented towards new procedures of intercepting, scanning and sorting big data,[84] new disease surveillance practices, made possible by the implementation of the algorithm to accrue and quantify troves of data, are now premised less on the idea of the *norm* or the *average* and alternatively are oriented towards anomaly detection within constantly expanding data sets. The infinite scope and capacity of algorithmic programming, to navigate vast datascapes, now means that even the most isolated or singular of aberrations or exceptional points can be taken into greater analysis and account. As Rouvroy asserts, the regulation of future risks will no longer be conveyed by exclusive references to the average. The aim and rationality within these intensifying systems of infinite data collection and production is to not miss any true positives embedded with digital data reservoirs, irrespective of the rate of false positives.[85]

V. Conclusion

The objective of this article has been to situate contemporary infectious disease surveillance practices to regulate pandemic risk within an era of "big data", and in doing so to consider and present how these significant algorithmic transformations alter and affect the rationalities of how such risks are understood, assessed and regulated. Critically, faced with a cascading of emergent and re-emergent infectious disease outbreaks and problematised by a historical legacy of an *incompleteness* of epidemic intelligence and surveillance data, practices of health surveillance and pandemic

[82] C Aradau and T Blanke, "Anomaly and the algorithmic subject of security" (2017) 3(1) European Journal of International Security 3.

[83] ibid.

[84] Krassman, supra, note 81.

[85] Rouvroy and Berns, supra, note 56, 9.

vigilance over the past two decades have sought with ever-evolving intensity to effectively harness the accelerated forecasting and prediction promises of big data to anticipate and report upon yet unforeseen pandemic episodes. As the analyses produced within this article has demonstrated, the implications and effects of this *digital* turn on strategies to assess and regulate the emergence of pandemic risks in a world of circulating pathogens and proliferating big data sources are far from trivial.

Centrally, the article has demonstrated how these broad and intensified efforts towards harnessing big data and the application of algorithmic techniques to enhance the real-time surveillance of infectious disease have produced critical transformations and shifts in rationalities and practices of infectious surveillance, manifested by the ascendancy of new digital surveillance devices which have progressively applied algorithmic processing logics to big data sets to produce new forms of expedited and digitised epidemic intelligence indicative of emergent or probable public health emergencies.

In transitioning from the era of the *avalanche of statistics* to the *deluge of big data*, these new surveillance systems represent new critical breaks away from practices of traditional health surveillance, moored within clinics and laboratories and informed exclusively by the statistics of the physical and pathogenic realms. The rationalities which underpin these risk technologies oscillate contrastingly towards processes of maximal accrual and analyses of open-source, disparate big data streams, extracted from emergent data ecologies and increasingly imparted by a range of new health security agents. Over the past two decades, the subsequent harnessing of big data to bolster and enhance the global forecasting and regulation of pathogenic threats has facilitated the emergence and rise of a new digitised surveillance regime which is increasingly automated, oriented towards the identification of exceptional public health emergencies, and indeed informed and guided by a new purveyor of epidemic intelligence – the digital algorithm.

Conceptually, in tracing the rise of big data-suffused surveillance systems for understanding and regulating pandemic risk, this article has argued that the steady unleashing of algorithms within the evolution of digital disease systems facilitates the ascendancy of new algorithmically-informed practices of infectious disease surveillance, indicative of new modes of *algorithmic government within health security and surveillance frameworks*. The rise of algorithmic governmentality produces new modes of algorithmic surveillance, as digital risk technologies for the regulation of pandemic illness, and these novel systems and practices diverge from the traditional parameters and function of governmentality developed by Foucault in several critical aspects, illustrated by transformative shifts from data scarcity to data excess; the emergence of novel data infrastructures and topographies, which alter the established surveillance roles of state institutions, clinicians and statisticians, and alternatively concretise new authorities of data monopolists, IT and data scientists in processes of disease surveillance; and finally, in charting novel divergences in algorithmic security logics which depart from a previous focus on governing emergent risks from the interface of the statistically-defined norm, towards digitally governing and identifying that which constitutes the exceptional or the aberrant within data sets.

And yet, despite such dramatic and far-reaching transformations in disease surveillance and pandemic regulation practices, critical accounts of ideologies of big

data, and the potentially dubious implications of the expanding authority of algorithmic government which underlie these expanding risk rationalities and practices in disease surveillance are largely absent within risk regulation and health security literature, save for the particularly illuminating work of Charlotte Heath-Kelly.[86] In the midst of these transformative risk assessment and surveillance practices, both critique and caution are merited.

As exemplified prominently by the dramatic rise and fall of Google Flu Trends, these increasingly big data-suffused surveillance systems have the potential for precise and timely reporting, evidenced by their operative reporting successes in the identification of SARS, H1N1 and Ebola, but also scope for spectacular failure, oversight and misclassifications. Moreover, it is the inherent and marked *opacity* of the algorithm as an increasingly salient security and regulatory actor which continues to serve as one of the most complex and exigent obstacles to understanding the design, function, gaze and unintended implications of these new infectious disease systems.[87] In seeking to digitally regulate pandemic via new rationalities of maximal data accrual and knowledge production via algorithmic processing, we cannot see from the outside, as Susanne Krasmann[88] writes, what kind of information, values or standards have been inscribed into these surveillance technologies. In a contemporary era fixated with the advanced identification of forthcoming security risks and imbued with the illuminative capacities of "big data", the NSA Revelations of 2013 demonstrated the enormous extent to which national governments and intelligence services have co-opted, enlisted and relied upon IT corporations and data monopolists including Microsoft, Google, Facebook and Twitter for the accrual, harvesting and transmitting of big data for the purpose of pre-emptive "security" calculations, estimations of risk and accelerated, dragnet surveillance. What currently remains less clear, however, within the politics of algorithmic surveillance practices, health security, and the pre-emption of pandemic risk, is how these same pervasive practices of risk analysis might come into direct conflict with human rights and privacy concerns during a potential public health emergency.

In seeking to address burgeoning conceptual, operational and technical gaps in assessing the opportunities to be gained for enhanced digital pandemic vigilance, and potentially negative implications of enhanced algorithmic disease surveillance in an era of big data, in 2018 the WHO released the *Guidelines on Ethical Issues in Public Health Surveillance*, in which it broadly acknowledged the significant shifting boundaries, expanse and scope of digital surveillance and sought to address issues of privacy, personal anonymity, as well as data validity and reliability of emerging digital disease practices. As such, within current health security frameworks which have demonstrated a growing and unbridled enthusiasm for emerging rationalities of risk analysis which are predicated upon intensified modes of algorithmic disease surveillance, and where the rise

[86] C Heath-Kelly, "Algorithmic autoimmunity in the NHS: Radicalisation and the clinic" (2017) 48(1) Security Dialogue 29.

[87] For an extended discussion on the ways in which algorithms increasingly mediate digital life, decision-making and augment or replace risk-assessment and decision-making by humans see further DB Mittlestadt et al, "The ethics of algorithms: Mapping the debate" (2016) Big Data & Society 1.

[88] Krasmann, supra, note 81, 21.

of advanced big data processing technologies and digital algorithms appears to have rapidly outpaced existent legislative and regulatory frameworks, the WHO's *Guidelines on Ethical Issues in Public Health Surveillance* must serve as critical catalyst towards engaged, constructive and sustained future efforts to evaluate, discuss and delineate the parameters of these emerging technologies and rationalities of risk, and, when necessary, to check and to **regulate** their extension.